Z. Wahrscheinlichkeitstheorie verw. Geb. 8, 259—297 (1967)

Strong Ratio Limits, R-recurrence and Mixing Properties of Discrete Parameter Markov Processes

F. PAPANGELOU *

Received January 13, 1967

Introduction

As is well-known a measure preserving transformation T on a finite measure space $(\Omega, \mathfrak{B}, \mu)$ is called mixing if

$$\lim_{n\to+\infty} \mu(\Omega)\,\mu(E\cap T^{-n}F) = \mu(E)\,\mu(F) \tag{1}$$

for all $E, F \in \mathfrak{B}$. One might think that a good analogue to (1) in the case of a σ-finite measure space is

$$\lim_{n\to+\infty} \frac{\mu(E\cap T^{-n}F)}{\mu(H\cap T^{-n}K)} = \frac{\mu(E)\,\mu(F)}{\mu(H)\,\mu(K)} \tag{2}$$

for all $E, F, H, K, \in \mathfrak{B}$ with finite and positive measure. However no transformation satisfying (2) for all such E, F, H, K exists [10, Th. 2], though it is perfectly legitimate to require the validity of (2) for all sets of finite and positive measure belonging to, say, a ring which is metrically dense in $\mathfrak{B}$, in the sense that for every $E \in \mathfrak{B}$ with finite measure and every $\varepsilon > 0$ there is an F in the ring with $\mu(E+F) < \varepsilon$ $(E+F = (E-F)\cup(F-E))$. K. KRICKEBERG [28], motivated by an example devised by E. HOPF, chose as such a ring the class of all "bounded" almost boundaryless sets, assuming of course that the space has a topology.

It turned out that the most interesting manifestations of this concept of mixing occur in the case of the shift transformation of Markov chains admitting an infinite stationary measure. In the case of recurrent chains for instance mixing is equivalent to the strong ratio limit property. There is more to it; when the measure is infinite there is room for variation. KRICKEBERG [28] introduced a concept of quasi-mixing by allowing the right-hand side of (2) to assume the form $\varphi(E,F)/\varphi(H,K)$, with some conditions on φ.

The main object of the present paper is to discuss mixing and quasi-mixing Markov chains. Only sections §1 and §7 are devoted to the abstract situation. We note however that we found it expedient, for reasons that will be explained below, to restrict the scope of KRICKEBERG's definition of quasi-mixing and accept only φ's of the form $\varphi(E,F) = \mu_1(E)\,\mu_2(F)$ where μ_1, μ_2 are measures.

In §1 we prove that, with this new definition, quasi-mixing in a space of finite measure coincides with mixing in the classical sense. This is not true if we permit more general φ' s. We also give a necessary condition for ergodicity which shows

* The present work was done while the author was holding a stipend of the Alexander von Humboldt foundation of W. Germany. A preliminary report was published in [32].

that there is an abundance of non-ergodic mixing transformations. In § 2 we specialize to Markov chains and show that the necessary and sufficient condition for quasi-mixing is essentially the strong ratio limit property of the transition probabilities (supported by another natural condition). We close the section with a discussion of these conditions. § 3 is devoted to what is perhaps the main theorem: Under a certain condition, satisfied in particular by all R-recurrent quasi-mixing chains, each of the measures μ_1, μ_2 is either absolutely continuous or singular relative to μ. The same is true of chains with independent increments, which are discussed in § 4. In § 5 we treat continuous-state-space processes with independent increments on locally compact groups, in the frame of STONE's paper [37]. Some discussion of rates of convergence, convergence norms etc. follows in § 6, and § 7 contains pathological examples of quasi-mixing transformations on the unit interval and the real line. A number of open problems are planted in the text.

The author is indebted to K. KRICKEBERG for many fruitful conversations and kindling questions as well as a lot of suggestions concerning the present paper. The "purity" of the measures μ_1, μ_2 relative to μ (see § 3 below) was conjectured by him.

§ 1. Mixing and Quasi-mixing Transformations

Let Ω be a completely regular topological space, $\mathfrak{B}$ the σ-field of its Borel sets, i. e. the σ-field generated by its open sets, and μ a σ-finite tight measure on $\mathfrak{B}$, which means that every $E \in \mathfrak{B}$ can be written as a union of countably many compact sets and a μ-negligible set. Let T be a $\mathfrak{B}$-measurable measure preserving transformation mapping all of Ω into itself and assume T is almost everywhere continuous, i.e. its discontinuity points form a μ-negligible set. If $\nu_1, \nu_2, \ldots, \nu_k$ are any measures on $\mathfrak{B}$ we denote by $\mathfrak{R}(\nu_1, \nu_2, \ldots, \nu_k)$ the field (not σ-field) of all sets in $\mathfrak{B}$ whose boundary is ν_i-negligible for every $i = 1, 2, \ldots, k$.

1.1. Definition. T is said to be *quasi-mixing* if there are two σ-finite tight measures μ_1, μ_2 on $\mathfrak{B}$, an increasing sequence of sets $H_k, k = 1, 2, \ldots$ in $\mathfrak{R}(\mu, \mu_1, \mu_2)$ and a sequence of positive numbers $\{\varrho_n\}$ such that

(i) $\mu(\Omega - \bigcup_k H_k) = 0\,, \quad \mu_1(\Omega - \bigcup_k H_k) = 0\,, \quad \mu_2(\Omega - \bigcup_k H_k) = 0\,,$

(ii) $0 < \mu(H_k) < +\infty\,, \quad 0 < \mu_1(H_k) < +\infty\,, \quad 0 < \mu_2(H_k) < +\infty$ for every k,

(iii) $\mu_1(E) > 0$, $\mu_2(E) > 0$ whenever E is in $\mathfrak{R}(\mu, \mu_1, \mu_2)$ and $\mu(E) > 0$, and finally,

(iv) for any two sets $E, F \in \mathfrak{R}(\mu, \mu_1, \mu_2)$ contained in some H_k the following is true

$$\lim_{n \to +\infty} \varrho_n \, \mu(E \cap T^{-n} F) = \mu_1(E) \, \mu_2(F)\,. \tag{3}$$

If $\{\varrho_n\}$, μ_1, μ_2 can be chosen so that $\mu_1 = \mu_2 = \mu$, T is termed *mixing*.

For the sequence $\{\varrho_n\}$ we shall use the expression "T is governed by $\{\varrho_n\}$". The role of the sequence $\{H_k\}$ is to introduce a notion of "boundedness" for subsets of Ω. Using the tightness of μ, μ_1, μ_2, the complete regularity of the space and conditions (i) and (ii) one can show that the class $\mathfrak{R}(\mu, \mu_1, \mu_2)$ is metrically dense in $\mathfrak{B}$ in the following sense: For every $A \in \mathfrak{B}$ and every $\varepsilon > 0$ there is $B \in \mathfrak{R}(\mu, \mu_1, \mu_2)$ such that $\mu(A + B) < \varepsilon$, $\mu_1(A + B) < \varepsilon$ and $\mu_2(A + B) < \varepsilon$. In [28] the class $\mathfrak{R}(\mu)$ of μ-almost boundaryless sets was used in place of $\mathfrak{R}(\mu, \mu_1, \mu_2)$ but the corresponding definition is far too strict and must be relaxed as above for otherwise the criterion of [28, p. 435] fails and the main examples in [28] satisfy only our weaker definition. In fact, if T is quasi-mixing and (3) holds for all μ-almost boundaryless sets contained in some H_k then μ_1 and μ_2 are absolutely continuous relative to μ, since μ, μ_1, μ_2 are tight and every μ-negligible compact set is μ-almost boundaryless. With the present definition the criterion of quasi-mixing referred to above remains true after a trivial modification and will be invoked on a number of occasions in the present paper.

Isomorphisms in the sense of [28] preserve mixing. Measure-preserving homeomorphisms preserve quasi-mixing. We note that (iii) is equivalent with the following: $\mu_1(E) > 0$, $\mu_2(E) > 0$ whenever E is open and $\mu(E) > 0$. One can further prove that (3) holds for every pair E, F such that E is μ_1-almost boundaryless, F is μ_2-almost boundaryless and $E \subset H_k$, $F \subset H_k$ for some k.

1.2. Theorem. *If T is a quasi-mixing transformation on a space of finite measure, then T is mixing in the classical sense.*

Proof. In [28] it is shown that $\{\varrho_n\}$ must be bounded and then inferred (with the aid of the metric density of $\mathfrak{R}(\mu, \mu_1, \mu_2)$ in $\mathfrak{B}$) that

$$\lim_{n \to +\infty} \mu(E \cap T^{-n} F) = \frac{\mu_1(E)\,\mu_2(F)}{\mu_1(\Omega)\,\mu_2(\Omega)}\,\mu(\Omega) \tag{4}$$

for *all* $E, F \in \mathfrak{B}$.

We shall first show that T is ergodic. Suppose there is an invariant set $E \in \mathfrak{B}$ with $0 < \mu(E) < \mu(\Omega)$. Then

$$0 = \lim_{n \to +\infty} \varrho_n\, \mu(E \cap T^{-n}(\Omega - E)) = \mu_1(E)\,\mu_2(\Omega - E)\,,$$

hence either $\mu_1(E) = 0$ or $\mu_2(\Omega - E) = 0$. If the first is true we have (since T is measure preserving and E is invariant) $\mu(E) = \mu(E \cap T^{-n}E) \to 0$ by (4), which is a contradiction. Similarly if $\mu_2(\Omega - E) = 0$.

Now the ergodicity of T implies the Cesàro convergence of $\mu(E \cap T^{-n}F)$ to $\mu(E)\,\mu(F)/\mu(\Omega)$, hence by (4)

$$\frac{\mu_1(E)\,\mu_2(F)}{\mu_1(\Omega)\,\mu_2(\Omega)}\,\mu(\Omega) = \frac{\mu(E)\,\mu(F)}{\mu(\Omega)}\,.$$

In a space of infinite measure a mixing transformation may fail to be ergodic:

1.3. Theorem. *If T is invertible and quasi-mixing and if*

$$\sum_{n=1}^{\infty} \frac{1}{\varrho_n} < +\infty\,,$$

then T is not ergodic.

Proof. By (3)

$$\lim_{n \to +\infty} \varrho_n\, \mu(H_1 \cap T^{-n} H_1) = \mu_1(H_1)\,\mu_2(H_1)\,.$$

This implies two things. First, since $1/\varrho_n \to 0$

(5) $$\lim_{n\to+\infty} \mu(H_1 \cap T^{-n} H_1) = 0 \quad \text{with} \quad \mu(H_1 \cap T^{-n} H_1) > 0 \quad \text{eventually.}$$

Second,

$$\mu(H_1 \cap T^{-n} H_1) < \tfrac{3}{2}\, \mu_1(H_1)\, \mu_2(H_1) \,.\, 1/\varrho_n$$

eventually, hence by our hypothesis

(6) $$\sum_{n=1}^{\infty} \mu(H_1 \cap T^{-n} H_1) < +\infty\,.$$

Suppose now T is ergodic. The space Ω has no atoms, because if A were an atom, we would have by the ergodicity of T $\Omega = \bigcup_{n=-\infty}^{+\infty} T^n A$ (almost), where the $T^n A$ are pairwise almost disjoint atoms in Ω with $\mu(T^n A) = \mu(A) > 0$ for every n. This however is impossible in view of (5). Thus Ω is non-atomic and T must be conservative (cf. [11, p. 85]). But then almost every point of H_1 is infinitely recurrent [11, p. 11], and this means

$$H_1 = \bigcup_{n=i}^{\infty} (H_1 \cap T^{-n} H_1) \quad \text{(almost) for every } i,$$

$$\mu(H_1) \leqq \sum_{n=i}^{\infty} \mu(H_1 \cap T^{-n} H_1) \quad \text{for every } i.$$

By (6) $\mu(H_1) = 0$, which is a contradiction.

Problem 1. What can one prove in the direction of sufficiency of

$$\sum_{n=i}^{\infty} \frac{1}{\varrho_n} = +\infty$$

for ergodicity? Does mixing plus conservativity imply ergodicity?

In this connection see the remark following Th. 2.1 below.

For each positive integer r we denote by $T^{(r)}$ the transformation $T^{(r)}(x_1, \ldots, x_r) = (Tx_1, \ldots, Tx_r)$ acting on the topological product space $\Omega^r = \Omega \times \cdots \times \Omega$ (r-fold). The product measure $\mu^{(r)} \mid \mathfrak{B}^{(r)}$, where $\mathfrak{B}^{(r)} = \mathfrak{B} \otimes \mathfrak{B} \otimes \cdots \otimes \mathfrak{B}$ (r-fold), can be extended to a tight measure on the σ-field $\mathfrak{B}^*$ of all Borel subsets of Ω^r [33]. It is easy to see that if $T^{(r)}$ is ergodic, then so is every $T^{(s)}$ with $s \leqq r$. Kakutani and Parry [17] define the *ergodic index* of T as the greatest positive integer r such that $T^{(r)}$ is ergodic; if $T^{(r)}$ is ergodic for every r then the ergodic index is by definition $+\infty$ and if T is not ergodic then the ergodic index is 0. As mentioned in [28], if T is quasi-mixing and governed by $\{\varrho_n\}$ then so is $T^{(r)}$ for every r and the latter is governed by $\{\varrho_n^r\}$. This and Th. 1.3 imply

1.4. Corollary. *If T is invertible and quasi-mixing, governed by $\{\varrho_n\}$, and if*

$$\sum_{n=1}^{\infty} \frac{1}{\varrho_n^r} < +\infty,$$

then the ergodic index of T is $< r$.

In connection with the proof of Th. 1.3 we refer the reader to Proposition 7.4.

1.5. Lemma. *If T is a quasi-mixing transformation governed by $\{\varrho_n\}$ and A a Borel set with $\mu(A) > 0$ and such that A and $T^{-1}A$ are members of $\mathfrak{K}(\mu, \mu_1, \mu_2)$ contained in some H_k then*

$$\lim_{n\to+\infty} \frac{\varrho_{n+1}}{\varrho_n} = \frac{\mu_1(T^{-1}A)}{\mu_1(A)} = \frac{\mu_2(A)}{\mu_2(T^{-1}A)} > 0 .$$

In fact

$$\frac{\varrho_{n+1}}{\varrho_n} = \frac{\varrho_{n+1}\,\mu(A\cap T^{-n}A)}{\varrho_n\,\mu(A\cap T^{-n}A)} = \frac{\varrho_{n+1}\,\mu(T^{-1}A\cap T^{-(n+1)}A)}{\varrho_n\,\mu(A\cap T^{-n}A)} \to \frac{\mu_1(T^{-1}A)\,\mu_2(A)}{\mu_1(A)\,\mu_2(A)} = \frac{\mu_1(T^{-1}A)}{\mu_1(A)}$$

and similarly

$$\frac{\varrho_{n+1}}{\varrho_n} = \frac{\varrho_{n+1}\,\mu(A\cap T^{-(n+1)}A)}{\varrho_n\,\mu(A\cap T^{-n}(T^{-1}A))} \to \frac{\mu_1(A)\,\mu_2(A)}{\mu_1(A)\,\mu_2(T^{-1}A)} = \frac{\mu_2(A)}{\mu_2(T^{-1}A)} .$$

Thus if our space contains a subset A satisfying the condition of the theorem then

$$\lim_{n\to+\infty} \frac{\varrho_{n+1}}{\varrho_n}$$

exists.

One can construct numerous mixing and quasi-mixing transformations by varying Hopf's original idea [16, pp. 66—67]. We do not pause to give such examples since a representation theorem of Krickeberg [28, § 2] shows that these transformations are "isomorphic" in a precise sense with Markov chains and enables us to construct them as "images" of the latter.

§ 2. Markov Chains

Let (p_{ij}) be the matrix of transition probabilities of a discrete parameter Markov chain with countable state space I. We assume

$$\sum_j p_{ij} = 1$$

for every i and denote the n-step transition probabilities by $p_{ij}^{(n)}$. We make the following two basic assumptions, *which will remain in force throughout the paper,* unless expressly stated otherwise:

Assumption I. The chain is irreducible and aperiodic.

Assumption II. The chain admits a non-trivial (finite or infinite) stationary measure.

The first assumption means that for any $i, j \in I$ there is $n_0(i, j)$ such that $p_{ij}^{(n)} > 0$ for all $n \geqq n_0(i, j)$. With the second assumption we postulate the existence of numbers λ_i, $i \in I$ with $0 < \lambda_i < +\infty$ for every $i \in I$ and

$$\lambda_j = \sum_i \lambda_i p_{ij}$$

for every $j \in I$. Algebraically $\{\lambda_i\}$ is a left eigenvector of the matrix (p_{ij}) corresponding to the eigenvalue 1. Such vectors will be called *positive stationary* *vectors* of the chain and will be referred to below as finite or infinite according as

$$\sum_i \lambda_i < +\infty \quad \text{or} \quad \sum_i \lambda_i = +\infty .$$

A stationary distribution is a positive stationary vector $\{\lambda_i\}$ with

$$\sum_i \lambda_i = 1 .$$

We are of course mainly interested in the case

$$\sum_i \lambda_i = +\infty .$$

In the sequel we shall choose one of the vectors $\{\lambda_i\}$ which will be called the stationary vector of the chain and will be at the basis of our considerations.

The measure-theoretic sample space of the Markov chain in question is constructed as follows. Let Ω be the set of all bilateral sequences of states

$$\Omega = \cdots \times I \times I \times I \times \cdots .$$

Elements of Ω will usually be denoted by x, with x_n designating the n-th coordinate ($n = 0, \pm 1, \pm 2, \ldots$). Let $\mathfrak{B}$ be the σ-field generated by the elementary rectangles, i.e. the sets which have the form

$$E = \{x\colon\ x_r = i_r,\ x_{r+1} = i_{r+1}, \ldots, x_s = i_s\}, \quad s \geqq r \tag{7}$$

and denote by μ the measure whose values on elementary rectangles such as E are given by

$$\mu(E) = \lambda_{i_r} p_{i_r i_{r+1}} p_{i_{r+1} i_{r+2}} \cdots p_{i_{s-1} i_s} . \tag{8}$$

On the σ-finite measure space $(\Omega, \mathfrak{B}, \mu)$ we introduce the "shift" transformation T which maps each point $x = (x_n)_n$ of Ω onto $Tx = (x_{n+1})_n$. Finally we topologize Ω with the product of the discrete topologies on the individual components. This topology derives from a metric which makes Ω into a Polish space, i.e. a complete separable metric space. $\mathfrak{B}$ is then the σ-field of Borel sets and every σ-finite measure on $\mathfrak{B}$ is tight. The transformation T is a homeomorphism of Ω onto itself and at the same time an invertible measure preserving transformation of $(\Omega, \mathfrak{B}, \mu)$.

An elementary rectangle such as E in (7) is clopen in the product topology, i.e. it has a void boundary. For each $i \in I$ we define $\Omega_i = \{x\colon\ x_0 = i\}$. Then

$$\Omega = \bigcup_i \Omega_i$$

where the Ω_i's are disjoint and $\mu(\Omega_i) = \lambda_i$.

The transformation T^{-1} "is" the shift of the reverse Markov chain, having transition probabilities

$$q_{ij} = \frac{\lambda_j}{\lambda_i} p_{ji} . \tag{9}$$

Note that the reverse chain depends on $\{\lambda_i\}$ and has the latter as a positive stationary vector.

A chain is called *reversible* if there are positive numbers $\{t_i\}$ such that

$$t_i p_{ij} = t_j p_{ji}$$

for all $i, j \in I$. $\{t_i\}$ is then a positive stationary vector for the chain and if we choose $\lambda_i = t_i$ (which we shall always do whenever we are dealing with a reversible chain) then the reverse chain is identical with the original one.

In applying Definition 1.1 to Markov chains we shall restrict somewhat its scope and confine ourselves to one particular sequence $\{H_k\}$: For each $m = 0, \pm 1, \pm 2, \ldots$ and each $i \in I$ we define $G_{m,i} = \{x\colon x_m = i\}$ and then suitably express each H_k as a finite union of such sets. To avoid making things look clumsier than they are we say that a subset of Ω is *bounded* if it is contained in a finite union of elementary rectangles. (3) is then expected to hold for all bounded members of $\mathfrak{K}(\mu, \mu_1, \mu_2)$. The chain will be called mixing or quasi-mixing if its shift transformation T is such. As shown in [28], in order to establish quasi-mixing it is sufficient to verify (3) for all rectangles of the form

$$\{x\colon x_{n_1} = i_1,\ x_{n_2} = i_2, \ldots, x_{n_k} = i_k\}, \quad n_1 < n_2 < \cdots < n_k .$$

2.1. Theorem. *The chain is mixing if and only if*

$$\lim_{n\to+\infty} \frac{p_{ij}^{(n+m)}}{p_{kh}^{(n)}} = \frac{\lambda_j}{\lambda_h} \tag{10}$$

for any states i, j, k, h and any integer m. If this is so and if i_0, j_0 are any states then the shift is governed by $\varrho_n = 1/p_{i_0 j_0}^{(n)}$.

Proof. (Cf. [28, end of § 3].) If $\mu_1 = \mu_2 = \mu$ and

$$E = \{x\colon x_r = i_r,\ x_{r+1} = i_{r+1}, \ldots, x_s = i_s\},$$

$$F = \{x\colon x_\nu = j_\nu,\ x_{\nu+1} = j_{\nu+1}, \ldots, x_t = j_t\}$$

then (3) can be written (note that $\nu + n > s$ eventually)

$$\lim_{n\to+\infty} \varrho_n \lambda_{i_r} p_{i_r i_{r+1}} \cdots p_{i_{s-1} i_s} p_{i_s j_\nu}^{(\nu+n)-s} p_{j_\nu j_{\nu+1}} \cdots p_{j_{t-1} j_t} = \lambda_{i_r} p_{i_r i_{r+1}} \cdots p_{i_{s-1} i_s} \cdot \lambda_{j_\nu} p_{j_\nu j_{\nu+1}} \cdots p_{j_{t-1} j_t}$$

i.e.

$$\lim_{n\to+\infty} \varrho_n p_{i_s j_\nu}^{(n+(\nu-s))} = \lambda_{j_\nu}.$$

The theorem follows.

A Markov chain constructed by Krengel [26, example 3.1] shows that mixing does not imply the validity of (3) for all (i.e. bounded as well as unbounded) members of $\mathfrak{K}(\mu, \mu_1, \mu_2)$ with finite measure.

It is interesting to compare Th. 2.1 with a condition for ergodicity. It is known [17] that T is ergodic if and only if the chain is recurrent, i.e.

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = +\infty .$$

If T is mixing this can be written

$$\sum_{n=1}^{\infty} \frac{1}{\varrho_n} = +\infty .$$

Thus we see that the necessary condition of Th. 1.3 is here sufficient as well.

Recall that an irreducible aperiodic chain is positive-recurrent if and only if it admits a stationary distribution.

2.2. Corollary (E. HOPF). *Every irreducible aperiodic Markov chain admitting a stationary distribution is mixing in the classical sense.*

In the null-recurrent case all one can say in general is that

$$\lim_{n\to+\infty} \frac{p_{ij} + p_{ij}^{(2)} + \cdots + p_{ij}^{(n)}}{p_{kh} + p_{kh}^{(2)} + \cdots + p_{kh}^{(n)}} = \frac{\lambda_j}{\lambda_h}.$$

If (10) holds, then we speak of a strong (or individual) ratio limit. Sufficient conditions for the validity of (10) in the null-recurrent case are given in [30] and [25]. It should be noted however that (10) is also satisfied by many transient chains which happen to admit a positive stationary vector.

Clearly (10) implies that the radius of convergence of the power series

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} x^n$$

is 1 for any i, j. D. VERE-JONES [38] has proved that for any irreducible and aperiodic Markov chain the radius of convergence R of the power series

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} x^n$$

is independent of i, j and that we either have

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} R^n = +\infty$$

for all i, j or

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} R^n < +\infty$$

for all i, j. In the former case he called the chain *R-recurrent* and in the latter *R-transient*. We always have $1 \leqq R < +\infty$ and it is known that

$$\lim_{n\to+\infty} \sqrt[n]{p_{ij}^{(n)}} = \frac{1}{R}$$

with $\sqrt[n]{p_{ii}^{(n)}} \leqq 1/R$ for every n [24, Th. 10]. The number $\gamma = 1/R$ is known as the *convergence norm* of the chain.

VERE-JONES further proved that if the chain is R-recurrent, then there exist positive numbers $\tau(i)$, $\pi(i)$ $(i \in I)$ such that

$$\gamma\tau(i) = \sum_j p_{ij}\tau(j) \quad \text{for every } i \in I, \tag{11}$$

$$\gamma\pi(j) = \sum_i \pi(i) p_{ij} \quad \text{for every } j \in I. \tag{12}$$

The vectors $\{\tau(i)\}$, $\{\pi(i)\}$ are the unique (to within a constant factor) non-negative solutions of (11) and (12) respectively and if we write

$$c_0 = \left(\sum_i \tau(i)\pi(i)\right)^{-1}$$

then

$$\lim_{n\to+\infty} p_{ij}^{(n)} R^n = c_0\tau(i)\pi(j) \quad \text{for all } i, j. \tag{13}$$

He called an R-recurrent chain *R-positive* or *R-null* according as $c_0 > 0$ or $= 0$. If the chain is R-positive we clearly have by (13)

$$\lim_{n\to+\infty} \frac{p_{ij}^{(n+m)}}{p_{kh}^{(n)}} = \gamma^m \frac{\tau(i)\,\pi(j)}{\tau(k)\,\pi(h)} \tag{14}$$

for any states i, j, k, h and any integer m. In the R-null case we have [38, Th. 3]

$$\lim_{n\to+\infty} \frac{p_{ij} R + p_{ij}^{(2)} R^2 + \cdots + p_{ij}^{(n)} R^n}{p_{kh} R + p_{kh}^{(2)} R^2 + \cdots + p_{kh}^{(2)} R^n} = \frac{\tau(i)\,\pi(j)}{\tau(k)\,\pi(h)}$$

but this does not imply (14). It implies however that if the limit in the left-hand side of (14) exists it must be equal to the right-hand side.

2.3. Definition (Pruitt [34]). An irreducible aperiodic Markov chain is said to have the *Strong Ratio Limit Property* (abbreviated SRLP) if there are positive numbers γ, $\tau(i)$, $\pi(i)$ $(i \in I)$ such that (14) is true for all states i, j, k, h and all integers m.

The sufficient conditions given by Orey [30] for the SRLP in the case of null-recurrent chains were generalized by Pruitt [34] to cover the R-null case. However, no universal criteria seem to be known for R-transient chains. We note that (14) implies that γ is the convergence norm of the chain, both in the R-recurrent and R-transient case.

Problem 2. Are there R-transient chains for which *all* individual ratio limits appearing in the left-hand side of (14) exist but do not have the form of the right-hand side?

Choose any state in I and denote it by 0. In the sequel it will play a special notational role. Recall that $\Omega_i = \{x\colon x_0 = i\}$.

2.4. Theorem. *The chain is quasi-mixing if and only if it has the SRLP and the numbers γ, $\tau(i)$, $\pi(i)$ $(i \in I)$ appearing in* (14) *satisfy equations* (11) *and* (12). *If this is so, and if E is the elementary rectangle in* (7) *then we can set*

$$\mu_1(E) = \gamma^{-s} \mu(E)\,\tau(i_s) = \gamma^{-s} \lambda_{i_r} p_{i_r i_{r+1}} \cdots p_{i_{s-1} i_s} \tau(i_s)\,, \tag{15}$$

$$\mu_2(E) = \gamma^r \mu(E) \frac{\pi(i_r)}{\lambda_{i_r}} = \gamma^r \pi(i_r)\, p_{i_r i_{r+1}} \cdots p_{i_{s-1} i_s} \tag{16}$$

and

$$\varrho_n = \frac{\tau(0)\,\pi(0)}{p_{00}^{(n)}}.$$

Conversely γ, $\tau(i)$, $\pi(i)$ $(i \in I)$ are defined in terms of $\{\varrho_n\}$, μ_1, μ_2 as follows:

$$\gamma = \lim_{n\to+\infty} \frac{\varrho_n}{\varrho_{n+1}},$$

$$\tau(i) = \frac{\mu_1(\Omega_i)}{\lambda_i},$$

$$\pi(i) = \mu_2(\Omega_i)\,.$$

Proof. The sufficiency of conditions (14), (11), (12) was established by Kricke-berg in [28] for his broader concept of quasi-mixing. Cf. the proof of Th. 2.1 above. To prove their necessity assume the shift is quasi-mixing. If A is any bounded almost boundaryless set of positive μ-measure, then by Th. 1.5

$$\lim_{n\to+\infty} \frac{\varrho_n}{\varrho_{n+1}} = \frac{\mu_1(A)}{\mu_1(T^{-1}A)} = \frac{\mu_2(T^{-1}A)}{\mu_2(A)} > 0 .$$

We define

$$\gamma = \lim_{n\to+\infty} \frac{\varrho_n}{\varrho_{n+1}}, \quad \tau(i) = \frac{\mu_1(\Omega_i)}{\lambda_i}, \quad \pi(i) = \mu_2(\Omega_i) \quad (i \in I)$$

and show that

$$\lim_{n\to+\infty} \varrho_n p_{ij}^{(n)} = \tau(i)\pi(j)$$

for every i, j. In fact

$$\varrho_n p_{ij}^{(n)} = \frac{\varrho_n \lambda_i p_{ij}^{(n)}}{\lambda_i} = \frac{\varrho_n \mu(\Omega_i \cap T^{-n}\Omega_j)}{\lambda_i} \to \frac{\mu_1(\Omega_i)\mu_2(\Omega_j)}{\lambda_i} = \tau(i)\pi(j) .$$

Finally

$$\frac{p_{ij}^{(n+m)}}{p_{kh}^{(n)}} = \frac{\varrho_n}{\varrho_{n+1}} \cdot \frac{\varrho_{n+1}}{\varrho_{n+2}} \cdots \frac{\varrho_{n+m-1}}{\varrho_{n+m}} \cdot \frac{\varrho_{n+m} p_{ij}^{(n+m)}}{\varrho_n p_{kh}^{(n)}}$$

which for $n \to +\infty$ converges to $\gamma^m \tau(i)\pi(j)/\tau(k)\pi(h)$. This establishes (14).

To prove (11) we note that

$$\Omega_i = \bigcup_j (\Omega_i \cap T^{-1}\Omega_j),$$

hence

$$\mu_1(\Omega_i) = \sum_j \mu_1(\Omega_i \cap T^{-1}\Omega_j).$$

This implies

$$\begin{aligned}
\gamma\tau(i) &= \gamma \frac{\mu_1(\Omega_i)}{\lambda_i} = \frac{\gamma}{\lambda_i} \sum_j \mu_1(\Omega_i \cap T^{-1}\Omega_j) \\
&= \frac{\gamma}{\lambda_i} \sum_j \frac{\mu_1(\Omega_i \cap T^{-1}\Omega_j)\mu_2(\Omega_0)}{\mu_2(\Omega_0)} \\
&= \frac{\gamma}{\lambda_i \pi(0)} \sum_j \left\{ \lim_{n\to+\infty} \varrho_n \mu(\Omega_i \cap T^{-1}\Omega_j \cap T^{-n}\Omega_0) \right\} \\
&= \frac{\gamma}{\lambda_i \pi(0)} \sum_j \left\{ \lim_{n\to+\infty} \varrho_n \lambda_i p_{ij} p_{j0}^{(n-1)} \right\} \\
&= \frac{\gamma}{\pi(0)} \sum_j \left\{ \lim_{n\to+\infty} \frac{\varrho_n}{\varrho_{n-1}} p_{ij} \varrho_{n-1} p_{j0}^{(n-1)} \right\} \\
&= \frac{\gamma}{\pi(0)} \sum_j \left\{ \frac{1}{\gamma} p_{ij} \tau(j)\pi(0) \right\} = \sum_j p_{ij}\tau(j) .
\end{aligned}$$

(12) can be established similarly. This concludes the proof.

We note that if a chain has the SRLP then so does the reverse chain (9), with the same γ and with $\tau^*(i) = \pi(i)/\lambda_i$, $\pi^*(i) = \lambda_i \tau(i)$. It follows that if the given chain is quasi-mixing, then so is the reverse one. It is known (see [30] and [34]) that a reversible and R-recurrent chain has the SRLP and is therefore quasi-mixing. It follows from the preceding remark that in this case

(17) $$\pi(i) = \lambda_i \tau(i) .$$

We shall now exhibit R-transient chains which have the SRLP but fail to satisfy one of the equations (12). Let f_n, $n = 1, 2, \ldots$, be a sequence of non-negative numbers with

$$\sum_{n=1}^{\infty} f_n \leqq 1$$

and define a Markov chain on the non-negative integers as follows

$$p_{01} = 1 - f_1, \quad p_{00} = f_1,$$

$$p_{i,i+1} = \frac{1 - f_1 - f_2 - \cdots - f_{i+1}}{1 - f_1 - f_2 - \cdots - f_i}, \quad p_{i0} = \frac{f_{i+1}}{1 - f_1 - f_2 - \cdots - f_i}.$$

The probability $f_{00}^{(n)}$ of first return to 0 at the n-th step (starting at 0) is equal to f_n (cf. [2, p. 42]). As usual we set $u_n = p_{00}^{(n)}$, $n = 1, 2, \ldots$, $u_0 = 1$, $f_0 = 0$. It is well-known that

$$u_n = f_1 u_{n-1} + f_2 u_{n-2} + \cdots + f_n u_0. \tag{18}$$

2.5. Proposition. *If in this model*

$$\lim_{n \to +\infty} \frac{u_{n+1}}{u_n} = \gamma$$

exists, then the model has the SRLP with

$$\tau(i) = \frac{\gamma^i - \sum_{\nu=1}^{i} f_\nu \gamma^{i-\nu}}{1 - \sum_{\nu=1}^{i} f_\nu}, \tag{19}$$

$$\pi(i) = \frac{1 - \sum_{\nu=1}^{i} f_\nu}{\gamma^i}. \tag{20}$$

Proof. Let us set $\alpha_i = 1 - f_1 - f_2 - \cdots - f_{i+1}$, $\alpha_{-1} = 1$, $\theta_i = \alpha_i/\alpha_{i-1}$. If i is any state other than 0 we obviously have, in the notation of [2, I § 9]

$$\begin{aligned} {}_0p_{i0}^{(n)} &= \theta_i \theta_{i+1} \cdots \theta_{i+n-2}(1 - \theta_{i+n-1}) = \frac{\alpha_{i+n-2}}{\alpha_{i-1}}\left(1 - \frac{\alpha_{i+n-1}}{\alpha_{i+n-2}}\right) = \\ &= \frac{\alpha_{i+n-2} - \alpha_{i+n-1}}{\alpha_{i-1}} = \frac{f_{i+n}}{\alpha_{i-1}} = \frac{{}_0p_{00}^{(i+n)}}{\alpha_{i-1}} \end{aligned}$$

hence

$$p_{i0}^{(n)} = {}_0p_{i0}^{(n)} + \sum_{\nu=1}^{n-1} {}_0p_{i0}^{(\nu)} p_{00}^{(n-\nu)} = \frac{{}_0p_{00}^{(i+n)}}{\alpha_{i-1}} + \sum_{\nu=1}^{n-1} \frac{{}_0p_{00}^{(i+\nu)}}{\alpha_{i-1}} p_{00}^{(n-\nu)}$$

by (18)

$$= \frac{1}{\alpha_{i-1}}\left[p_{00}^{(i+n)} - \sum_{\mu=1}^{i} {}_0p_{00}^{(\mu)} p_{00}^{(i+n-\mu)}\right]$$

$$= \frac{1}{\alpha_{i-1}}\left[u_{i+n} - \sum_{\mu=1}^{i} f_\mu u_{i+n-\mu}\right]$$

which implies

$$\lim_{n \to +\infty} \frac{p_{i0}^{(n)}}{p_{00}^{(n)}} = \lim_{n \to +\infty} \frac{1}{\alpha_{i-1}}\left[\frac{u_{i+n}}{u_n} - \sum_{\mu=1}^{i} f_\mu \frac{u_{n+i-\mu}}{u_n}\right] = \frac{1}{\alpha_{i-1}}\left[\gamma^i - \sum_{\mu=1}^{i} f_\mu \gamma^{i-\mu}\right]. \tag{21}$$

Note that $p_{i0}^{(n)} \geqq p_{i0} p_{00}^{(n-1)}$ hence

$$\lim_{n \to +\infty} \frac{p_{i0}^{(n)}}{p_{00}^{(m)}} \geqq p_{i0} \gamma^{-1} > 0.$$

Now let j be $\neq 0$ and i arbitrary.

$$p_{ij}^{(n)} = {}_0p_{ij}^{(n)} + \sum_{\nu=1}^{n-1} p_{i0}^{(\nu)}\, {}_0p_{j0}^{(n-\nu)}.$$

When $n > j$

$$p_{ij}^{(n)} = p_{i0}^{(n-j)}\, {}_0p_{0j}^{(j)} = p_{i0}^{(n-j)} \alpha_{j-1}. \tag{22}$$

Clearly (21) and (22) are trivially true when $i = 0$ or $j = 0$. Hence given any four states i, j, k, h

$$\frac{p_{ij}^{(n+m)}}{p_{kh}^{(n)}} = \frac{p_{i0}^{(n+m-j)}\,\alpha_{j-1}}{p_{k0}^{(n-h)}\,\alpha_{h-1}} = \frac{\dfrac{p_{i0}^{(n+m-j)}}{p_{00}^{(n+m-j)}} \cdot \dfrac{p_{00}^{(n+m-j)}}{p_{00}^{(n)}} \cdot \alpha_{j-1}}{\dfrac{p_{k0}^{(n-h)}}{p_{00}^{(n-h)}} \cdot \dfrac{p_{00}^{(n-h)}}{p_{00}^{(n)}} \cdot \alpha_{h-1}}.$$

When $n \to +\infty$ this converges to

$$\frac{\dfrac{1}{\alpha_{i-1}}\left[\gamma^i - \sum_{\mu=1}^{i} f_\mu \gamma^{i-\mu}\right] \cdot \gamma^{m-j} \cdot \alpha_{j-1}}{\dfrac{1}{\alpha_{k-1}}\left[\gamma^k - \sum_{\mu=1}^{k} f_\mu \gamma^{k-\mu}\right] \cdot \gamma^{-h} \cdot \alpha_{h-1}} = \gamma^m \frac{\dfrac{1}{\alpha_{i-1}}\left[\gamma^i - \sum_{\mu=1}^{i} f_\mu \gamma^{i-\mu}\right]}{\dfrac{1}{\alpha_{k-1}}\left[\gamma^k - \sum_{\mu=1}^{k} f_\mu \gamma^{k-\mu}\right]} \cdot \frac{\dfrac{\alpha_{j-1}}{\gamma^j}}{\dfrac{\alpha_{h-1}}{\gamma^h}}$$

which completes the proof of the proposition.

We can easily verify that

$$\gamma\, \tau(i) = \sum_j p_{ij}\, \tau(j)$$

for every i. Similarly, for every $j \neq 0$,

$$\gamma\, \pi(j) = \sum_i \pi(i)\, p_{ij}.$$

However the equation

$$\gamma\, \pi(0) = \sum_i \pi(i)\, p_{i0}$$

is satisfied if and only if

$$\gamma = \sum_{i=0}^{\infty} \frac{\alpha_{i-1}}{\gamma^i}(1 - \theta_i) = \sum_{i=0}^{\infty} \frac{f_{i+1}}{\gamma^i}$$

i.e. if and only if

$$\sum_{i=1}^{\infty} f_i R^i = 1.$$

But this is easily seen to be necessary and sufficient for R-recurrence. Thus

2.6. Proposition. *If a chain of the present model has the SRLP then* (11) *and* (12) *hold if and only if the chain is R-recurrent.*

For concrete examples consider any Markov chain having the SRLP, let f_n be the probability of first return to 0 at the n-th step, starting at 0, and use these numbers in the above construction. The model is not wholly satisfactory since only one of the equalities fails in the R-transient case and (what is worse) it admits a positive stationary vector if and only if it is recurrent (see [5, p. 544]). This leaves us with the following task:

Problem 3. Construct a chain having the SRLP and admitting a positive stationary vector, but such that (11) and/or (12) do not hold.

It is of course readily seen that the SRLP implies

$$\gamma\,\tau(i) \geqq \sum_j p_{ij}\,\tau(j) \quad \text{for every } i, \tag{23}$$

$$\gamma\,\pi(j) \geqq \sum_i \pi(i)\,p_{ij} \quad \text{for every } j, \tag{24}$$

and we cannot resist the temptation to mention in parenthesis the following

2.7. Proposition. *The inequalities* (23) *with* $\tau(i) \geqq 0$ $(i \in I)$ *imply*

$$\inf_i \tau(i) = 0$$

when $0 < \gamma < 1$ *and*

$$\inf_i \Big[\tau(i) - \sum_j p_{ij}\,\tau(j)\Big] = 0$$

when $\gamma = 1$. *(Assumptions* I *and* II *are not needed here.)*

Proof. The case $0 < \gamma < 1$ is trivial since iteration of (23) yields

$$\gamma^n\,\tau(i) \geqq \sum_j p_{ij}^{(n)}\,\tau(j) \geqq \Big(\sum_j p_{ij}^{(n)}\Big)\cdot \inf_j \tau(j) = \inf_j \tau(j)\,.$$

Let now $\gamma = 1$ and set

$$\varepsilon = \inf_i \Big[\tau(i) - \sum_j p_{ij}\,\tau(j)\Big].$$

We prove that for any n

$$\sum_j p_{ij}^{(n)}\,\tau(j) \geqq \sum_j p_{ij}^{(n+1)}\,\tau(j) + \varepsilon\,.$$

In fact

$$\begin{aligned}\sum_j p_{ij}^{(n)}\,\tau(j) &\geqq \sum_j p_{ij}^{(n)}\Big[\sum_h p_{jh}\,\tau(h) + \varepsilon\Big]\\ &= \sum_h \tau(h)\Big(\sum_j p_{ij}^{(n)}\,p_{jh}\Big) + \varepsilon \sum_j p_{ij}^{(n)}\\ &= \sum_h \tau(h)\,p_{ih}^{(n+1)} + \varepsilon\,.\end{aligned}$$

An induction argument now leads to

$$\tau(i) \geqq \sum_j p_{ij}^{(n)}\,\tau(j) + n\,\varepsilon$$

for every n, whence $\varepsilon = 0$.

For the dual inequalities (24) (with $0 < \gamma \leqq 1$) one can say the following: (i) If there is a j_0 such that

$$\sum_i p_{ij_0} = +\infty$$

then

$$\inf_j \pi(j) = 0\,.$$

(ii) If there is a positive stationary vector $\{\lambda_i\}$, then

$$\inf_j \frac{\pi(j)}{\lambda_j} = 0$$

in the case $\gamma < 1$ while

$$\inf_j \frac{\pi(j) - \sum_i \pi(i) p_{ij}}{\lambda_j} = 0$$

in the case $\gamma = 1$. (i) is trivial and to prove (ii) we apply Prop. 2.7 to the reverse chain (9) which satisfies the inequalities

$$\gamma \frac{\pi(j)}{\lambda_j} \geqq \sum_i q_{ji} \frac{\pi(i)}{\lambda_i}.$$

2.8. Theorem. *If a chain is quasi-mixing and transient then* $\mu_1(\Omega) = +\infty$, $\mu_2(\Omega) = +\infty$. *(In the recurrent case* $\mu_1 = \mu_2 = \mu$*).*

Proof. If $\gamma < 1$ then (24) implies

$$\sum_j \pi(j) = +\infty$$

and if $\{\lambda_i\}$ is a stationary vector then (23) implies

$$\sum_i \lambda_i \tau(i) = +\infty.$$

In fact

$$\gamma \sum_j \pi(j) \geqq \sum_j \sum_i \pi(i) p_{ij} = \sum_i \pi(i) \sum_j p_{ij} = \sum_i \pi(i)$$

and

$$\gamma \sum_i \lambda_i \tau(i) \geqq \sum_j \Big(\sum_i \lambda_i p_{ij}\Big) \tau(j) = \sum_j \lambda_j \tau(j).$$

In particular if the chain is quasi-mixing then

$$\mu_1(\Omega) = \sum_i \mu_1(\Omega_i) = \sum_i \lambda_i \tau(i) + \infty$$

and

$$\mu_2(\Omega) = \sum_j \mu_2(\Omega_j) = \sum_j \pi(j) = +\infty.$$

If $\gamma = 1$ then $\mu_2(\Omega) < +\infty$ if and only if the chain is positive-recurrent. In fact $\{\pi(i)\}$ is a positive stationary vector and

$$\sum_j \pi(j) < +\infty$$

means that the chain admits a stationary distribution. Similarly $\mu_1(\Omega) < +\infty$ if and only if the chain is positive-recurrent; for, by what has just been shown, $\mu_1(\Omega) < +\infty$ if and only if the reverse chain (9) is positive-recurrent and this is obviously true if and only if the original chain is positive-recurrent.

§ 3. The Main Theorem

In the present section we prove that in the case of an R-recurrent quasi-mixing Markov chain each of the measures μ_1, μ_2 is "pure" (i.e. either singular or absolutely continuous) relative to μ. Here is a slightly more general result.

3.1. Theorem. *If* γ *and* $\tau(i)$, $i \in I$ *are any positive numbers satisfying equations* (11) *then formula* (15) *determines a measure* μ_1 *on* $\mathfrak{B}$, *and if* $\{\tau(i)\}$ *is the unique* *(to within a constant factor) positive solution of* (11) *then the measure* μ_1 *is either absolutely continuous or singular relative to* μ. *If, in addition to the above hypotheses,* $\gamma = 1$ *then* $\mu_1 = c\mu$ *where c is a positive constant.*

Proof. Equations (11) ensure the validity of KOLMOGOROV's compatibility conditions for the set function μ_1 which can thus be extended to a measure on $\mathfrak{B}$. Decompose this measure μ_1 into absolutely continuous and singular parts relative to μ,

$$\mu_1(A) = \int_A f(x)\,\mu(dx) + \theta(A) \qquad (A \in \mathfrak{B}). \tag{25}$$

In the space Ω we define a net $\{\mathfrak{M}_n\}$ as follows. For each fixed $n \geqq 0$ $\mathfrak{M}_n$ consists of all rectangles of the form

$$E_{i_{-n}\cdots i_0 \cdots i_n} = \{x\colon\ x_{-n} = i_{-n}, \ldots, x_0 = i_0, \ldots, x_n = i_n\}.$$

It is clear that

$$\bigcup_{n=0}^{\infty} \mathfrak{M}_n$$

generates $\mathfrak{B}$, hence by a well-known theorem of DE LA VALLÉE-POUSSIN [35, pp. 152—156] the derivative of μ_1 with respect to μ on the above net is f. More specifically, if for each $x = (\ldots, x_{-1}, x_0, x_1, \ldots)$ we define

$$f_n(x) = \begin{cases} \dfrac{\mu_1(E_{x_{-n}\cdots x_0\cdots x_n})}{\mu(E_{x_{-n}\cdots x_0\cdots x_n})} & \text{if} \quad \mu(E_{x_{-n}\cdots x_0\cdots x_n}) > 0 \\ 0 & \text{if} \quad \mu(E_{x_{-n}\cdots x_0\cdots x_n}) = 0, \end{cases}$$

then

$$f(x) = \lim_{n\to+\infty} f_n(x) \qquad \text{a.e.}$$

(cf. [6, Chap. VII, § 8, especially p. 346 or pp. 611—612]).

Now from (15) we obviously have $f_n(x) = \gamma^{-n}\tau(x_n)$ a.e., therefore

$$f(x) = \lim_{n\to+\infty} \gamma^{-n}\tau(x_n) \qquad \text{a.e.} \tag{26}$$

and we assume that $f(x)$ is this limit or 0 according as the limit exists or not. From (26)

$$f(Tx) = \gamma f(x) \qquad \text{a.e.}$$

Further f is integrable on each Ω_i since

$$\int_{\Omega_i} f(x)\,\mu(dx) \leqq \lim_{n\to+\infty} \int_{\Omega_i} f_n(x)\,\mu(dx) = \lim_{n\to+\infty} \mu_1(\Omega_i) = \mu_1(\Omega_i).$$

At this point we employ a technique of [1]. From (26) it follows that given any k the function f is measurable relative to the σ-field generated by $x_k, x_{k+1}, \ldots$. Hence by the Markov property if $n_1 < n_2 < \cdots < n_r$ we have for the conditional expectations:

$$E(f \mid x_{n_1} = i_1, \ldots, x_{n_r} = i_r) = E(f \mid x_{n_r} = i_r)$$

Therefore

$$\gamma^n E(f \mid x_0 = i_0, \ldots, x_n = i_n) = E(f \mid x_0 = i_n) \tag{27}$$

since

$$E(\gamma^n f \mid x_0 = i_0, \ldots, x_n = i_n) = E(\gamma^n f \mid x_n = i_n)$$
$$= E(f \circ T^n \mid (T^n x)_0 = i_n) = E(f \mid x_0 = i_n).$$

In particular (27) implies

$$\gamma E(f \mid x_0 = i, x_1 = j) = E(f \mid x_0 = j),$$

hence if we define $\sigma(i) = E(f \mid x_0 = i)$ we get

$$\gamma \sigma(i) = \gamma E(f \mid x_0 = i)$$
$$= \gamma \sum_j E(f \mid x_0 = i,\ x_1 = j)\, p_{ij} = \sum_j E(f \mid x_0 = j)\, p_{ij} = \sum_j p_{ij}\, \sigma(j),$$

which shows that $\{\sigma(i)\}$ is a non-negative solution of (11). There are now two possibilities:

Case (i). $\sigma(i_0) = 0$ for some i_0.

Then $\sigma(j) = 0$ for all j, since

$$\gamma \sigma(i) = \sum_j p_{ij}\, \sigma(j) \quad \text{implies} \quad \gamma^n \sigma(i_0) = \sum_j p_{i_0 j}^{(n)}\, \sigma(j)$$

for all n and by the irreducibility of the chain for any j there is n with $p_{i_0 j}^{(n)} > 0$.

In this case f must be 0 a.e. and hence μ_1 is singular relative to μ.

Case (ii). $\sigma(i) > 0$ for all i.

In this case there is an $\alpha > 0$ such that $\tau(i) = \alpha \sigma(i)$ for every i, i.e.

$$\tau(i) = \alpha E(f \mid x_0 = i).$$

This and (27) now imply that given any states $i_0, i_1, \ldots, i_n$

$$\gamma^{-n} \tau(i_n) = \gamma^{-n} \alpha E(f \mid x_0 = i_n) = \alpha E(f \mid x_0 = i_0, \ldots, x_n = i_n)$$

which can be written

$$\gamma^{-n} \tau(x_n) = E(\alpha f \mid x_0, x_1, \ldots, x_n).$$

Thus we see that $\gamma^{-n} \tau(x_n)$ is a uniformly integrable martingale on each Ω_i (normalized into a probability space) converging to $\alpha f(x)$ a.e. [27, Chap. IV, Th. 3.6]. Comparing with (26) we see that $\alpha = 1$ and the uniform integrability together with (25) imply

$$\mu_1(\Omega_i) \leqq \int_{\Omega_i} f(x)\, \mu(dx) = \lim_{n \to +\infty} \int_{\Omega_i} \gamma^{-n} \tau(x_n)\, \mu(dx) = \mu_1(\Omega_i)$$

i. e.

$$\mu_1(\Omega_i) = \int_{\Omega_i} f(x)\, \mu(dx)$$

whence

$$\theta(\Omega_i) = 0, \quad \theta(\Omega) = \sum_i \theta(\Omega_i) = 0.$$

Finally if $\gamma = 1$, then the constant $v(i) = 1$ is a solution of (11), therefore $\tau(i) = c$. By (15) $\mu_1 = c\mu$.

The first corollary below is actually a consequence of (26).

3.2. Corollary. *Retaining hypothesis* (11) *only, let* $\{\lambda_i'\}$ *be another positive stationary vector for the chain, denote by* μ' *the invariant measure induced by* $\{\lambda_i'\}$ *on*

$\mathfrak{B}$ *and define* μ_i' *by* (15), *writing* λ_i' *in place of* λ_i. *Then* μ_1' *is singular relative to* μ' *if and only if* μ_1 *is singular relative to* μ.

Proof. Let $\mathfrak{B}^0$ be the σ-subfield of $\mathfrak{B}$ generated by $x_0, x_1, x_2, \ldots$. The restrictions of μ and μ' to $\mathfrak{B}^0$ are absolutely continuous relative to each other because they are proportional on each Ω_i. Now μ_1 is singular relative to μ if and only if

$$\lim_{n\to+\infty} \gamma^{-n}\,\tau(x_n) = 0$$

μ-almost everywhere and the latter is true if and only if

$$\lim_{n\to+\infty} \gamma^{-n}\,\tau(x_n) = 0$$

μ'-almost everywhere, since the exceptional set belongs to $\mathfrak{B}^0$.

3.3. Corollary. *If* γ *and* $\pi(i)$, $i \in I$ *are positive numbers satisfying equations* (12) *and if* $\{\pi(i)\}$ *is the unique (to within a constant factor) positive solution of this equation, then the measure* μ_2 *determined by* (16) *is either absolutely continuous or singular relative to* μ. *If in addition* $\gamma = 1$, *then* $\mu_2 = c\,\mu$.

Proof. All we have to do is to apply Th. 3.1 to the reverse chain (9). If we set $\tau^*(i) = \pi(i)/\lambda_i$ then it follows from our hypothesis and (9) that $\{\tau^*(i)\}$ is the unique (to within a constant factor) solution of

$$\gamma\,\tau^*(i) = \sum_j q_{ij}\,\tau^*(j) \qquad (i \in I)\,.$$

Let $(\Omega, \mathfrak{B}, \mu^*)$ be the sample space of the new chain and define μ_1^* by (15), writing $\tau^*(i)$ in place of $\tau(i)$ and q's in place of p's. By Th. 3.1 μ_1^* is either absolutely continuous or singular relative to μ^*. Now if S is the transformation of Ω which maps (x_n) onto (x_{-n}), then $\mu_2 = \mu_1^* \circ S$, $\mu = \mu^* \circ S$. We conclude that μ_2 is either absolutely continuous or singular relative to μ.

We note that the analogue of (26) for μ_2 is

$$\frac{d\mu_2}{d\mu}(x) = \lim_{n\to+\infty} \gamma^{-n}\,\frac{\pi(x_{-n})}{\lambda_{x_{-n}}} \text{ a.e.}$$

3.4. Corollary. *If a chain is quasi-mixing and if* $\{\tau(i)\}$ ($\{\pi(i)\}$) *is the unique positive solution of* (11) ((12)), *then the measure* μ_1 (μ_2) *is either absolutely continuous or singular relative to* μ. *If both* $\{\tau(i)\}$ *and* $\{\pi(i)\}$ *are unique and if* $\gamma = 1$, *then we have mixing. The conclusions are true in particular if the chain is R-recurrent.*

Problem 4. Is μ_1 (μ_2) "pure" even if $\{\tau(i)\}$ ($\{\pi(i)\}$) is not the unique positive solution of (11) ((12))?

Problem 5. Is it true that if $\gamma < 1$ then μ_1, μ_2 are singular relative to μ?

This seems unlikely. There are however two classes of Markov chains for which it holds: R-positive chains (Th. 3.5) and chains with independent increments (§ 4 and § 5).

Problem 6. If the answer to problem 5 is negative, find a necessary and sufficient condition for singularity.

Problem 7. Again, if the answer to problem 5 is negative, can μ_1 be absolutely continuous while μ_2 is singular relative to μ?

This is impossible in the case of reversible chains (Th. 3.6).

3.5. Theorem. *If a chain is R-positive and if $\gamma < 1$ (i.e. $R > 1$) then μ_1 and μ_2 are singular relative to μ.*

Proof. Consider the restrictions of μ, μ_1, μ_2 to the finite measure space Ω_0. If

$$A_n = \{x\colon\ x_0 = 0, \quad x_n = 0\},$$
$$B_n = \{x\colon\ x_{-n} = 0, \quad x_0 = 0\}$$

then $\mu(A_n) = \mu(B_n) = \lambda_0\, p_{00}^{(n)} \to 0$ since the chain is transient, while

$$\mu_1(A_n) = \gamma^{-n}\lambda_0\, p_{00}^{(n)}\,\tau(0) = \lambda_0\,\tau(0)\, p_{00}^{(n)}\, R^n \to \lambda_0\,\tau(0)\, c_0\,\tau(0)\,\pi(0) > 0$$

and

$$\mu_2(B_n) = \gamma^{-n}\pi(0)\, p_{00}^{(n)} = \pi(0)\, p_{00}^{(n)}\, R^n \to \pi(0)\, c_0\,\tau(0)\,\pi(0) > 0$$

i.e. μ_1, μ_2 cannot be absolutely continuous relative to μ.

If an R-recurrent chain is reversible then the transformation S which maps each point (x_n) of Ω onto (x_{-n}) satisfies $\mu = \mu \circ S$ and $\mu_2 = \mu_1 \circ S$ (by (17)). Hence:

3.6. Theorem. *In the case of a reversible R-recurrent chain the measures μ_1, μ_2 are either both absolutely continuous or both singular relative to μ.*

§ 4. Chains with Independent Increments

Throughout the present section we shall consider Markov chains with independent increments (i.e. spatially homogeneous random walks) on the integers. In other words we assume that the state space I is the set of all integers and that

$$p_{ij} = p_{0,\,j-i}\,.$$

We also postulate, as always, irreducibility and aperiodicity. Notice that a chain with independent increments admits the positive stationary vector $\lambda_i = 1$ ($i \in I$).

Let

$$g(s) = \sum_{i=-\infty}^{+\infty} p_{0i}\, s^i \qquad (0 < s < +\infty)\,.$$

Then

4.1 ([3], [19]; see [23] for a polished treatment). *Every irreducible aperiodic random walk has the SRLP with $\gamma = g(s_0)$, $\tau(i) = s_0^i$, $\pi(i) = s_0^{-i}$, where s_0 is the unique positive number with*

$$0 < g(s_0) = \inf_s g(s) \leqq 1\,.$$

The case $R = 1$ was treated by CHUNG and ERDÖS. KEMENY later reduced the general case to this one by introducing the random walk

$$q_{0i} = \frac{p_{0i}\, s_0^i}{g(s_0)}\,. \tag{28}$$

It is interesting that the formula behind (28) is

$$q_{0i} = R\,\frac{\tau(i)}{\tau(0)}\, p_{0i}$$

which is exactly the transformation employed by VERE-JONES [38] (see the proof of his Theorem II) to reduce an R-recurrent chain to a recurrent (i. e. 1-recurrent) one.

4.2. Theorem. *Every irreducible and aperiodic random walk on the integers is quasi-mixing. If $\gamma = 1$ we have mixing; if $\gamma < 1$, both μ_1 and μ_2 are singular relative to μ.*

There is no overlapping with Th. 3.5, since a random walk is never R-positive. In fact if it is R-recurrent then

$$\sum_i \tau(i)\,\pi(i) = \sum_i s_0^i s_0^{-i} = \sum_i 1 = +\infty .$$

Before proving this theorem we note a few auxiliary facts about the generating function g. Obviously $g(1) = 1$. The domain of finiteness Δ of g is either $\{1\}$ or an interval in $(0, +\infty)$ containing 1. The end-points s_1, s_2 of Δ may or may not belong to Δ and

$$0 \leqq s_1 \leqq 1 \leqq s_2 \leqq +\infty .$$

Consider the derivative

$$g'(s) = \sum_{i=-\infty}^{+\infty} i\,p_{0i}\,s^{i-1}$$

of g.

Case (i). If $\Delta = \{1\}$ then we put

$$g'(1) = \sum_{i=-\infty}^{+\infty} i\,p_{0i}$$

if the series is well-defined, i.e. either the positive or negative "side" is finite. The only case where $g'(1)$ is undefined is when

$$\sum_{i>0} i\,p_{0i} = +\infty \quad \text{and} \quad \sum_{i<0} i\,p_{0i} = -\infty .$$

Case (ii). If $s_1 < s_2$ then $g'(s)$ exists and is equal to

$$\sum_{i=-\infty}^{+\infty} i\,p_{0i}\,s^{i-1}$$

for every $s \in \Delta$, with the understanding that if $s_1 \in \Delta$ then $g'(s_1)$ is one-sided and may be either finite or $-\infty$ while if $s_2 \in \Delta$ then $g'(s_2)$ is again one-sided and may be finite or $+\infty$.

The functions g and g' are continuous in Δ, g is strictly convex, while g' is strictly increasing. Let X be a random variable having the distribution of the increments of the chain, i.e. $\text{Prob}\{X = i\} = p_{0i}$. If $g'(1)$ is well-defined, then it is equal to the "expectation" of X, which we denote by $E(X)$ and allow to be $+\infty$ or $-\infty$.

4.3. Lemma. *If the minimum s_0 of g is $\neq 1$ and if $g'(1)$ is finite, then*

$$\frac{[s_0]^{g'(1)}}{g(s_0)} < 1 .$$

Proof. Assume $s_0 < 1$. Then $[s_0, 1] \subset \Delta$. Let us set $\lambda = g'(1)$ for convenience. Since g' is strictly increasing and since s_0 is the minimum of g we must have

$g'(s) > 0$ for every $s \in (s_0, 1]$; in particular $\lambda = g'(1) > 0$. Now define the functions

$$\varphi(s) = \frac{s g'(s)}{g(s)}, \quad \psi(s) = \frac{s^\lambda}{g(s)}, \qquad s_0 \leqq s \leqq 1 .$$

First we show (see also [19]) that φ is strictly increasing. In fact

$$\varphi'(s) = \frac{g(s)[s g''(s) + g'(s)] - s g'(s)^2}{g(s)^2} \tag{29}$$

where

$$\begin{aligned} s g''(s) + g'(s) &= s \sum_i i(i-1)\, p_{0i}\, s^{i-2} + \sum_i i\, p_{0i}\, s^{i-1} \\ &= \sum_i i^2 p_{0i}\, s^{i-1} - \sum_i i\, p_{0i}\, s^{i-1} + \sum_i i\, p_{0i}\, s^{i-1} \\ &= \sum_i i^2 p_{0i}\, s^{i-1} . \end{aligned}$$

Thus the numerator in (29) is equal to

$$\Big(\sum_i p_{0i}\, s^i\Big)\Big(\sum_i i^2 p_{0i}\, s^{i-1}\Big) - s\Big(\sum_i i\, p_{0i}\, s^{i-1}\Big)^2$$

which is positive, since by the Schwarz inequality

$$\begin{aligned} s\Big(\sum_i i\, p_{0i}\, s^{i-1}\Big)^2 &= s\Big(\sum_i \sqrt{p_{0i} s^{i-1}} \cdot i \sqrt{p_{0i} s^{i-1}}\Big)^2 < \\ &< s\Big(\sum_i p_{0i} s^{i-1}\Big)\Big(\sum_i i^2 p_{0i}\, s^{i-1}\Big) \\ &= \Big(\sum_i p_{0i}\, s^i\Big)\Big(\sum_i i^2 p_{0i}\, s^{i-1}\Big) \end{aligned}$$

Since φ is strictly increasing, $\varphi(s) < \varphi(1)$ i.e.

$$\frac{s g'(s)}{g(s)} < \lambda \quad \text{for every} \quad s \in [s_0, 1) \tag{30}$$

hence $g(s) \cdot \lambda - s g'(s) > 0$ for every $s \in [s_0, 1)$. This now implies

$$\psi'(s) = \frac{s^{\lambda-1}}{g(s)^2}\,[g(s)\,\lambda - s g'(s)] > 0 \quad \text{for every} \quad s \in [s_0, 1)$$

which means that ψ is strictly increasing in $[s_0, 1]$; in particular

$$\psi(s_0) < \psi(1)$$

i. e.

$$\frac{s^\lambda}{g(s_0)} < 1 .$$

The case $s_0 > 1$ can be treated similarly; alternatively we can consider the reverse random walk with $q_{0i} = p_{0,-i}$ $(= p_{i0})$.

Proof of Theorem 4.2. Quasi-mixing follows from Theorem 2.4, since equations (11) and (12) all collapse to

$$g(s_0) = \sum_{i=-\infty}^{+\infty} p_{0i}\, s_0^i$$

which is true!

If $\gamma = 1$ then by 4.1 $s_0 = 1$, $\tau(i) = \pi(i) = 1$ for every i and (15), (16) show that $\mu_1 = \mu_2 = \mu$.

If $\gamma < 1$ then $s_0 \neq 1$. By (26) and 4.1 if $d\mu_1/d\mu$ denotes the Radon-Nikodym derivative of μ_1 with respect to μ, then for almost every $x = (x_n)$ in Ω

$$\frac{d\mu_1}{d\mu}(x) = \lim_{n\to+\infty} \gamma^{-n}\tau(x_n) = \lim_{n\to+\infty} \frac{s_0^{x_n}}{g(s_0)^n} = \lim_{n\to+\infty}\left(\frac{s_0^{x_n/n}}{g(s_0)}\right)^n . \tag{31}$$

By the strong law of large numbers x_n/n converges to $E(X) = g'(1)$ for almost every x. (This is also true in the cases $E(X) = +\infty$ and $E(X) = -\infty$, by an elementary truncation argument.) If $g'(1)$ is finite, the lemma and (31) show that $d\mu_1/d\mu = 0$ a.e. If $g'(1)$ is infinite we distinguish two cases: In the case $s_0 < 1$ $g'(1)$ can only be $+\infty$, since g' is strictly increasing. Thus $x_n/n \to +\infty$ a.e. and since $s_0 < 1$ we have by (31) $d\mu_1/d\mu = 0$ a.e. In the case $s_0 > 1$ $g'(1)$ can only be $-\infty$, hence $x_n/n \to -\infty$ a.e , $d\mu_1/d\mu = 0$ a.e.

We conclude that in all cases $d\mu_1/d\mu = 0$ a.e. hence μ_1 is singular relative to μ. Singularity of μ_2 now follows if we consider the reverse random walk. This completes the proof.

Note that neither R-recurrence nor uniqueness of solutions for (11) and (12) were needed. In fact this theorem covers R-transient random walks as well. In this connection the following theorem is not without interest. Recall that $g'(s_0)$ is undefined only when $s_0 = 1$ and

$$\sum_{i>0} i\, p_{0i} = +\infty\,, \quad \sum_{i<0} i\, p_{0i} = -\infty\,.$$

4.4. Theorem. *If $g'(s_0)$ is well-defined then the random walk is R-null-recurrent or R-transient according as $g'(s_0) = 0$ or $\neq 0$.*

Proof. The random walk is R-recurrent if and only if

$$\sum_{n=1}^{\infty} p_{00}^{(n)} R^n = +\infty$$

i. e.

$$\sum_{n=1}^{\infty} \frac{p_{00}^{(n)}}{g(s_0)^n} = +\infty\,.$$

Now consider the random walk whose transition probabilities are given by (28). It is easy to see that

$$q_{00}^{(n)} = \frac{p_{00}^{(n)}}{g(s_0)^n}$$

therefore

$$\sum_{n=1}^{\infty} \frac{p_{00}^{(n)}}{g(s_0)^n} = +\infty$$

is equivalent to

$$\sum_{n=1}^{\infty} q_{00}^{(n)} = +\infty\,.$$

In other words the original random walk is R-recurrent if and only if the new one is recurrent and this is the case if and only if the expectation $E(Y)$ of the increment Y of the new chain is 0 (see [36, T 1, p. 33]; the cases $E(Y) = +\infty$ or $-\infty$ are transient). Now

$$E(Y) = \sum_{i=-\infty}^{+\infty} i\, q_{0i} = \sum_{i=-\infty}^{+\infty} i\, \frac{p_{0i} s_0^i}{g(s_0)} = \frac{s_0\, g'(s_0)}{g(s_0)}\,.$$

The proposition follows if we recall a remark made earlier to the effect that no random walk is R-positive.

It is interesting to note that by (30), which was proved under the assumptions "$s_0 < 1$, $g'(1)$ finite" we have

$$0 \leqq E(Y) < E(X).$$

This is also true when $g'(1)$ is infinite. If $s_0 > 1$ then

$$E(X) < E(Y) \leqq 0.$$

In the case $s_0 = 1$ the two random walks coincide. Thus Y is in general closer to being centred.

In connection with Th. 4.4 we refer the reader to [4] for criteria of recurrence (= 1-recurrence) in the case $s_0 = 1$,

$$\sum_{i>0} i\, p_{0i} = +\infty, \quad \sum_{i<0} i\, p_{0i} = -\infty.$$

4.5. Corollary. *If s_0 is an interior point of Δ (in particular if X has finitely many possible values) then the random walk is R-recurrent and $g'(1) = E(X)$ is well-defined. Further $\gamma = 1$ or < 1 according as $E(X) = 0$ or $\neq 0$, i.e. the random walk is mixing if and only if the increment X is centred.*

In fact s_0 is the minimum of g, hence $g'(s_0) = 0$. If X has finitely many possible values then $\Delta = (0, +\infty)$. Compare [36, p. 51]. If X has infinitely many possible values, then the random walk can be mixing, even though X may not be centred.

Examples. *(a)* The symmetric random walk in three dimensions is 1-transient.

(b) An R-transient random walk with $R > 1$:

$$g(s) = c\left(\sum_{i=1}^{\infty} \frac{1}{i^3\, 2^{i+2}} \frac{1}{\zeta(2)} s^{-i} + 1 + \sum_{i=1}^{\infty} \frac{1}{2^i} s^i\right)$$

where

$$\zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2}$$

and c is a normalizing constant. Here $\Delta = [\frac{1}{2}, 2)$, the minimum occurs at $s_0 = \frac{1}{2}$ and $g'(\frac{1}{2}) > 0$.

(c) Let

$$g(s) = c\left(\sum_{i=1}^{\infty} \frac{1}{2^i} s^{-i} + \frac{1}{\zeta(2)} \sum_{i=1}^{\infty} \frac{1}{i^3} s^i\right).$$

Here $\Delta = (\frac{1}{2}, 1]$, the minimum occurs at 1 and $g'(1) < 0$. Thus $\gamma = 1$ (i.e. we have mixing) but $g'(1) \neq 0$ (i.e. the increment is not centred). This is a 1-transient chain.

(d)

$$g(s) = \frac{1}{3\,\zeta(3)}\left(\sum_{i=1}^{\infty} \frac{1}{i^3} \frac{1}{s^i} + \sum_{i=1}^{\infty} \frac{2}{i^3} s^i\right).$$

Here $\Delta = \{1\}$ and $E(X) = g'(1)$ exists, is finite and > 0.

$$\left(\zeta(3) \text{ denotes } \sum_{k=1}^{\infty} \frac{1}{k^3}\right)$$

This chain is 1-transient.

§ 5. Continuous State Space

In the present section we shall discuss a Markov process with discrete time parameter and independent increments on a locally compact Abelian group G. We denote its transition probabilities by $p(\xi, A)$ ($\xi \in G$, $A \in \mathfrak{F}$; here $\mathfrak{F}$ is the σ-field of Borel subsets of G, i.e. the σ-field generated by the open sets of G).

The strong ratio limit theorem 4.1 of Chung-Erdös and Kemeny was generalized by Ornstein [31] and C. Stone [37]. We present here part of Stone's general result. Following Stone we assume that

(i) G is compactly generated, i.e. there is a compact subset F of G such that the smallest group containing F is G [15, p. 35, (5.12)]. This implies that G is σ-compact.

(ii) $p(0, \cdot)$ is a Radon measure on $\mathfrak{F}$.

(iii) The closed semi-group generated by the support of $p(0, \cdot)$ is G itself.

The hypothesis of independent increments means

$$p(\xi, A) = p(0, A - \xi)$$

where $A - \xi = \{y - \xi \colon y \in A\}$. The Haar measure λ on G is stationary for the process.

Denote by $\mathfrak{S}$ the collection of all continuous homomorphisms of G into R (the real line) and define

$$g(s) = \int_G e^{s(\xi)} p(0, d\xi), \qquad s \in \mathfrak{S}.$$

Let $\mathfrak{A}^*$ be the class of all relatively compact λ-almost boundaryless subsets of G and $\mathfrak{A}$ the class of all $A \in \mathfrak{A}^*$ with $\lambda(A) > 0$. Then

5.1 (Stone [37]). *Under hypotheses* (i), (ii), (iii) *there is a unique* $s_0 \in \mathfrak{S}$ *such that* $0 < g(s_0) = \inf_{s \in \mathfrak{S}} g(s) \leqq 1$ *and if* $A \in \mathfrak{A}$, $B \in \mathfrak{A}$ *then*

$$\lim_{n \to +\infty} \frac{p^{(n+m)}(\xi, A)}{p^{(n)}(\eta, B)} = g(s_0)^m e^{s_0(\xi - \eta)} \frac{\int_A e^{-s_0(z)} \lambda(dz)}{\int_B e^{-s_0(z)} \lambda(dz)}$$

uniformly with respect to ξ *and* η *in compact sets.*

We shall employ this result to establish quasi-mixing. As before let

$$\Omega = \cdots \times G \times G \times G \times \cdots$$

be our sample space, topologized with the product topology, and denote by $\mathfrak{R}$ the semi-ring of all rectangles of the form

$$(32) \qquad E = \{x \in \Omega \colon x_r \in A_r,\ x_{r+1} \in A_{r+1}, \ldots, x_k \in A_k\}, \qquad k \geqq r$$

with $A_i \in \mathfrak{F}$, $i = r, r+1, \ldots, k$. The set function

$$\mu(E) = \int_{A_r} \lambda(dx_r) \int_{A_{r+1}} p(x_r, dx_{r+1}) \int_{A_{r+2}} \cdots \int_{A_k} p(x_{k-1}, dx_k)$$

can be uniquely extended to a σ-finite measure on the product σ-field $\mathfrak{B}_0$ (compare with [29, p. 78]) which is invariant under the shift transformation T. In the present context (see § 1) we wish to extend μ to the σ-field $\mathfrak{B}$ of all Borel subsets of Ω relative to the product topology. $\mathfrak{B}$ is in general larger than $\mathfrak{B}_0$ and it will be shown in another note [33] that there is a unique tight measure on $\mathfrak{B}$ (to be denoted again by μ) extending $\mu \mid \mathfrak{R}$.

If there is a countable basis of neighborhoods of 0 in G, then $\mathfrak{B}_0 = \mathfrak{B}$ and G and Ω are Polish spaces. Hence every σ-finite measure on Ω is tight.

A subset of Ω will be called *bounded* if it is contained in a finite union of rectangles such as E in (32), having relatively compact components $A_r, A_{r+1}, \ldots, A_k$. With this notion of boundedness we have.

5.2. Theorem. *Under hypotheses* (i), (ii), (iii) *the random walk is quasi-mixing. The values of* μ_1, μ_2 *on rectangles such as E in* (32) *are given by*

$$\mu_1(E) = g(s_0)^{-k} \int_{A_r} \lambda(dx_r) \int_{A_{r+1}} p(x_r, dx_{r+1}) \int_{A_{r+2}} \cdots \int_{A_k} p(x_{k-1}, dx_k)\, e^{s_0(x_k)}, \tag{33}$$

$$\mu_2(E) = g(s_0)^{r} \int_{A_r} e^{-s_0(x_r)} \lambda(dx_r) \int_{A_{r+1}} p(x_r, dx_{r+1}) \int_{A_{r+2}} \cdots \int_{A_k} p(x_{k-1}, dx_k) \tag{34}$$

with $s_0(\xi)$, $\xi \in G$ *as defined in* 5.1. *If* $s_0(\xi) \equiv 0$, *we have mixing.*

In the course of the proof we shall need a lemma:

5.3. Lemma. *Let m be any σ-finite measure on* $\mathfrak{F}$. *If* $f(x_1, x_2, \ldots, x_k)$ *is a non-negative function on* $G \times G \times \cdots \times G$ *(k-fold) which is measurable relative to the product σ-field* $\mathfrak{F} \otimes \mathfrak{F} \otimes \cdots \otimes \mathfrak{F}$ *then*

$$\int_G m(dx_1) \int_G p(x_1, dx_2) \int_G \cdots \int_G p(x_{k-1}, dx_k)\, f(x_1, x_2, \ldots, x_k) =$$
$$= \int_G \int_G \cdots \int_G f(y_1, y_1 + y_2, \ldots, y_1 + y_2 + \cdots + y_k)\, m(dy_1)\, p(0, dy_2) \cdots p(0, dy_k).$$

Proof. We transform the innermost integral by means of the transformation $y_k = x_k - x_{k-1}$ (fixed x_{k-1}) and then proceed setting successively

$$y_{k-1} = x_{k-1} - x_{k-2}, \ldots, y_2 = x_2 - x_1,\ y_1 = x_1.$$

5.4. Corollary. *If m is any σ-finite measure on* $\mathfrak{F}$ *and if* $A_1, A_2, \ldots, A_k \in \mathfrak{F}$ *then*

$$\int_{A_1} m(dx_1) \int_{A_2} p(x_1, dx_2) \int_{A_3} \cdots \int_{A_k} p(x_{k-1}, dx_k) =$$
$$= \int\int_{G^{k-1}}\int \cdots \left[\int_{Z_{y_2 \cdots y_k}} m(dy_1) \right] p(0, dy_2) \cdots p(0, dy_k)$$

where

$$Z_{y_2 \cdots y_k} = A_1 \cap [A_2 \cap (\cdots (A_{k-1} \cap (A_k - y_k) - y_{k-1}) \cdots) - y_2].$$

In fact if $f(x_1, x_2, \ldots, x_k)$ in 5.3 is the indicator of $A_1 \times A_2 \times \cdots \times A_k$ in G^k then $f(y_1, y_1 + y_2, \ldots, y_1 + y_2 + \cdots + y_k)$ is the indicator of

$$Z_{y_2 \cdots y_k} \times \underbrace{G \times G \times \cdots \times G}_{k-1 \text{ times}}.$$

Note also the formula

$$\int_G \lambda(dx) \int_G p(x, dy)\, h(y) = \int_G \lambda(dy)\, h(y)$$

which holds for any non-negative measurable function h on G and which implies (if we insert indicators):

$$\int_G \lambda(dx) \int_A p(x, dy)\, h(y) = \int_A \lambda(dy)\, h(y) \qquad (A \in \mathfrak{F}). \tag{35}$$

Proof of Theorem 5.2. We first show that μ_1, μ_2 as given by (33), (34) are well-defined, i.e. satisfy the necessary compatibility conditions.

If $A_k = G$ in (33), then the innermost integral is equal to

$$\begin{aligned}\int_G p(x_{k-1}, dx_k)\, e^{s_0(x_k)} &= \int_G p(0, dt)\, e^{s_0(t+x_{k-1})}\\ &= e^{s_0(x_{k-1})} \int_G p(0, dt)\, e^{s_0(t)} = e^{s_0(x_{k-1})}\, g(s_0)\end{aligned}$$

(recall that $s_0(t + x_{k-1}) = s_0(t) + s_0(x_{k-1})$ since s_0 is a homomorphism).

If $A_r = G$, then by (35) with

$$h(y) = \int_{A_{r+2}} p(y, dx_{r+2}) \int_{A_{r+3}} \cdots \int_{A_k} p(x_{k-1}, dx_k)\, e^{s_0(x_k)}$$

we have

$$\begin{aligned}&g(s_0)^{-k} \int_G \lambda(dx_r) \int_{A_{r+1}} p(x_r, dx_{r+1}) \int_{A_{r+2}} \cdots \int_{A_k} p(x_{k-1}, dx_k)\, e^{s_0(x_k)}\\ &= g(s_0)^{-k} \int_{A_{r+1}} \lambda(dx_{r+1}) \int_{A_{r+2}} \cdots \int_{A_k} p(x_{k-1}, dx_k)\, e^{s_0(x_k)}.\end{aligned}$$

Thus the definition of μ_1 is consistent.

In the case of μ_2, if $A_k = G$ then

$$\int_{A_k} p(x_{k-1}, dx_k) = 1$$

and everything is trivial. If on the other hand $A_r = G$ then we use 5.4:

$$\mu_2(E) = g(s_0)^r \int\!\!\int_{G^{k-r}}\!\! \cdots \int \left[\int_{Z_{y_{r+1}\cdots y_k}} e^{-s_0(\xi)}\, \lambda(d\xi)\right] p(0, dy_{r+1}) \cdots p(0, dy_k) \tag{36}$$

where

$$\begin{aligned}Z_{y_{r+1}\cdots y_k} &= A_r \cap [A_{r+1} \cap (A_{r+2} \cap (\cdots) - y_{r+2}) - y_{r+1}]\\ &= A_{r+1} \cap (A_{r+2} \cap (\cdots) - y_{r+2}) - y_{r+1} \equiv Z^*_{y_{r+2}\cdots y_k} - y_{r+1}.\end{aligned}$$

Since λ is translation invariant we have

$$\int_{Z_{y_{r+1}\cdots y_k}} e^{-s_0(\xi)}\, \lambda(d\xi) = \int_{Z^*_{y_{r+2}\cdots y_k} - y_{r+1}} e^{-s_0(\xi)}\, \lambda(d\xi) = \int_{Z^*_{y_{r+2}\cdots y_k}} e^{-s_0(\xi - y_{r+1})}\, \lambda(d\xi) = e^{s_0(y_{r+1})} \int_{Z^*_{y_{r+2}\cdots y_k}} e^{-s_0(\xi)}\, \lambda(d\xi).$$

Substituting in (36)

$$\begin{aligned}\mu_2(E) &= g(s_0)^r \left(\int_G e^{s_0(y_{r+1})}\, p(0, dy_{r+1})\right) \times\\ &\quad \times \left(\int\!\!\int_{G^{k-r-1}}\!\! \cdots \int \left[\int_{Z^*_{y_{r+2}\cdots y_k}} e^{-s_0(\xi)}\, \lambda(d\xi)\right] p(0, dy_{r+2}) \cdots p(0, dy_k)\right)\\ &= g(s_0)^{r+1} \int_{A_{r+1}} e^{-s_0(x_{r+1})}\, \lambda(dx_{r+1}) \int_{A_{r+2}} p(x_{r+1}, dx_{r+2}) \int_{A_{r+3}} \cdots \int_{A_k} p(x_{k-1}, dx_k)\end{aligned}$$

by the definition of $g(s)$ and 5.4 again. The compatibility conditions have thus been verified. The theorem on p. 78 of [29] (modified so as to cover σ-finite measures) now implies that μ_1, μ_2 can be uniquely extended to two measures on the product σ-field $\mathfrak{B}_0$. As with μ, one can show (see [33]) that there are unique tight extensions of μ_1, μ_2 to the σ-field $\mathfrak{B}$ of all Borel subsets of Ω.

To establish quasi-mixing it is sufficient to establish (3) for all rectangles E whose components $A_r, A_{r+1}, \ldots, A_k$ are in $\mathfrak{A}^*$ [28, § 1]. Such rectangles belong to $\mathfrak{R}(\mu, \mu_1, \mu_2)$; see 5.3 and 5.4.

By 5.1 if we fix $B_0 \in \mathfrak{A}$ and set

$$\varrho_n = \frac{\int\limits_{B_0} e^{-s_0(z)} \lambda(dz)}{p^{(n)}(0, B_0)}$$

then for every $A \in \mathfrak{A}$

$$\lim_{n \to +\infty} \varrho_n p^{(n+m)}(\xi, A) = g(s_0)^m e^{s_0(\xi)} \int\limits_A e^{-s_0(z)} \lambda(dz) \tag{37}$$

uniformly with respect to ξ if ξ is restricted to a compact set. This is easily seen to be true for every $A \in \mathfrak{A}^*$ (if $\lambda(A) = 0$ write $A = B - (B - A)$ for some $B \in \mathfrak{A}$).

Let now E be given by (32) and

$$F = \{x\colon\ x_\tau \in B_\tau,\ x_{\tau+1} \in B_{\tau+1}, \ldots, x_\nu \in B_\nu\}, \qquad \nu \geqq \tau$$

where all the A_i's and B_j's are in $\mathfrak{A}^*$. Then

$$\varrho_n \mu(E \cap T^{-n} F) = \varrho_n \mu(\{x\colon\ x_r \in A_r, \ldots, x_k \in A_k,\ x_{\tau+n} \in B_\tau, \ldots, x_{\nu+n} \in B_\nu\}).$$

When $n \to +\infty$ $\tau + n$ is eventually greater than k, hence

$$\begin{aligned}\varrho_n \mu(E \cap T^{-n} F) &= \varrho_n \int\limits_{A_r} \lambda(dx_r) \int\limits_{A_{r+1}} p(x_r, dx_{r+1}) \int\limits_{A_{r+2}} \cdots \\ &\quad \cdots \int\limits_{A_k} p(x_{k-1}, dx_k) \int\limits_{B_\tau} p^{(\tau+n-k)}(x_k, dx_{\tau+n}) \int\limits_{B_{\tau+1}} \cdots \int\limits_{B_\nu} p(x_{\nu+n-1}, dx_{\nu+n}) \\ &= \varrho_n \int\limits_{A_r} \lambda(dx_r) \int\limits_{A_{r+1}} p(x_r, dx_{r+1}) \int\limits_{A_{r+2}} \cdots \\ &\quad \cdots \int\limits_{A_k} p(x_{k-1}, dx_k) \Big[\int\int\limits_{G^{\nu-\tau}} \cdots \int p^{(\tau+n-k)}(x_k, Z_{y_{\tau+1} \cdots y_\nu}) \, p(0, dy_{\tau+1}) \cdots p(0, dy_\nu) \Big]\end{aligned}$$

by Corollary 5.4, where $Z_{y_{\tau+1} \cdots y_\nu} = B_\tau \cap (\cdots (B_\nu - y_\nu) - \cdots)$ is in $\mathfrak{A}^*$. Now by (37)

$$\lim_{n \to +\infty} \varrho_n p^{(\tau+n-k)}(x_k, B_\tau) = g(s_0)^{\tau-k} e^{s_0(x_k)} \int\limits_{B_\tau} e^{-s_0(z)} \lambda(dz) \tag{38}$$

uniformly in $x_k \in A_k$; in particular there is a constant M and a positive integer n_0 such that

$$\varrho_n p^{(\tau+n-k)}(x_k, B_\tau) \leqq M \quad \text{for all} \quad x_k \in A_k \quad \text{and all} \quad n \geqq n_0$$

(the right-hand side of (38) is bounded on A_k as a function of x_k). Since $Z_{y_{\tau+1} \cdots y_\nu} \subset B_\tau$ for any $y_{\tau+1}, \ldots, y_\nu$, we have

$$\varrho_n p^{(\tau+n-k)}(x_k, Z_{y_{\tau+1} \cdots y_\nu}) \leqq M \quad \text{for all} \quad x_k \in A_k, \quad y_{\tau+1} \in G, \ldots, y_\nu \in G \quad \text{and all} \quad n \geqq n_0 \tag{39}$$

while by (37)

$$\lim_{n\to+\infty} \varrho_n\, p^{(\tau+n-k)}(x_k, Z_{y_{\tau+1}\cdots y_\nu}) = g(s_0)^{\tau-k} e^{s_0(x_k)} \int_{Z_{y_{\tau+1}\cdots y_\nu}} e^{-s_0(z)} \lambda(dz).$$

By Lebesgue's dominated convergence theorem

$$\lim_{n\to+\infty} \int\int_{G^{\nu-\tau}}\cdots\int \varrho_n\, p^{(\tau+n-k)}(x_k, Z_{y_{\tau+1}\cdots y_\nu})\, p(0, dy_{\tau+1})\cdots p(0, dy_\nu)$$
$$= \int\int_{G^{\nu-\tau}}\cdots\int \left[g(s_0)^{\tau-k} e^{s_0(x_k)} \int_{Z_{y_{\tau+1}\cdots y_\nu}} e^{-s_0(z)} \lambda(dz)\right] p(0, dy_{\tau+1})\cdots p(0, dy_\nu)$$

which we write for simplicity

$$\lim_{n\to+\infty} F_n(x_k) = F(x_k) \quad \text{for all} \quad x_k \in A_k.$$

By (39), for every $x_k \in A_k$ and every $n \geqq n_0$

$$F_n(x_k) \leqq \int\int_{G^{\nu-\tau}}\cdots\int M\, p(0, dy_{\tau+1})\cdots p(0, dy_\nu) = M$$

and Lebesgue's theorem again yields

$$\lim_{n\to+\infty} \int_{A_k} p(x_{k-1}, dx_k)\, F_n(x_k) = \int_{A_k} p(x_{k-1}, dx_k)\, F(x_k).$$

Iterating this argument and bearing in mind that $\lambda(A_r)$ is finite we arrive at

$$\lim_{n\to+\infty} \varrho_n\, \mu(E \cap T^{-n} F) = \int_{A_r} \lambda(dx_r) \int_{A_{r+1}} p(x_r, dx_{r+1}) \int_{A_{r+2}}\cdots$$
$$\cdots \int_{A_k} p(x_{k-1}, dx_k) \int\int_{G^{\nu-\tau}}\cdots\int \left[g(s_0)^{\tau-k} e^{s_0(x_k)} \int_{Z_{y_{\tau+1}\cdots y_\nu}} e^{-s_0(z)} \lambda(dz)\right] p(0, dy_{\tau+1})\cdots p(0, dy_\nu).$$

By Corollary 5.4 this is equal to

$$g(s_0)^{-k} \int_{A_r} \lambda(dx_r) \int_{A_{r+1}} p(x_r, dx_{r+1}) \int_{A_{r+2}}\cdots \int_{A_k} p(x_{k-1}, dx_k)\, e^{s_0(x_k)} \times$$
$$\times\, g(s_0)^{\tau} \int_{B_\tau} e^{-s_0(x_\tau)} \lambda(dx_\tau) \int_{B_{\tau+1}} p(x_\tau, dx_{\tau+1}) \int_{B_{\tau+2}}\cdots \int_{B_\nu} p(x_{\nu-1}, dx_\nu) = \mu_1(E)\, \mu_2(F).$$

Finally, if $s_0(\xi) \equiv 0$, then obviously $\mu_1 = \mu_2 = \mu$ and the proof is complete.

Let us specialize to the case $G = R$ (the real line). Ω is then a Polish space, $\mathfrak{B} = \mathfrak{B}_0$ and $\mathfrak{S}$ can be identified with R. The generating function $g(s)$ becomes

$$g(s) = \int_R e^{s\xi}\, p(0, d\xi), \qquad s \in R.$$

As in the discrete case the domain of finiteness of $g(s)$ is an interval Δ containing 0 which may be infinite or degenerate. The function g and its derivative g' have properties entirely analogous to those of the corresponding generating function in § 4, with 0 here playing the role of 1 there. As before, let s_0 be the minimum of $g(s)$ (see 5.1).

5.5. Lemma. *If $s_0 \neq 0$ and if $g'(0)$ is finite, then*

$$\frac{e^{s_0 g'(0)}}{g(s_0)} < 1.$$

Proof. Suppose $s_0 < 0$; then $[s_0, 0] \subset \Delta$. Let $\lambda_0 = g'(0)$ and define

$$\varphi(s) = \frac{g'(s)}{g(s)}, \quad \psi(s) = \frac{e^{s\lambda_0}}{g(s)}, \quad s_0 \leqq s \leqq 0.$$

Then $\varphi(s)$ is strictly increasing in $[s_0, 0]$ since

$$\varphi'(s) = \frac{g(s)g''(s) - g'(s)^2}{g(s)^2}$$

and by the Schwarz inequality

$$g'(s)^2 = \left(\int \xi e^{s\xi} p(0, d\xi)\right)^2 = \left(\int \sqrt{e^{s\xi}} \cdot \xi \sqrt{e^{s\xi}} p(0, d\xi)\right)^2 < \\ < \left(\int e^{s\xi} p(0, d\xi)\right)\left(\int \xi^2 e^{s\xi} p(0, d\xi)\right) = g(s)g''(s)$$

for every $s \in [s_0, 0)$. We infer $\varphi(s) < \varphi(0)$ for every $s \in [s_0, 0)$, i.e. $g'(s)/g(s) < \lambda_0$ whence $g(s)\lambda_0 - g'(s) > 0$ for every $s \in [s_0, 0)$. But then

$$\psi'(s) = \frac{e^{s\lambda_0}[g(s)\lambda_0 - g'(s)]}{g(s)^2} > 0 \quad \text{in} \quad [s_0, 0)$$

which implies that $\psi(s)$ is strictly increasing and in particular $\psi(s_0) < \psi(0) = 1$, i.e.

$$\frac{e^{s_0\lambda_0}}{g(s_0)} < 1 .$$

The case $s_0 > 0$ can be treated similarly.

5.6. Theorem. *If $s_0 = 0$ (i.e. $g(s_0) = 1$) the random walk is mixing. If $s_0 \neq 0$ (i.e. $g(s_0) < 1$) the random walk is quasi-mixing and the measures μ_1, μ_2 are singular relative to μ.*

Proof. Suppose $s_0 \neq 0$. As in the discrete case $g'(0)$ is well-defined (though possibly infinite) and by the strong law of large numbers

$$\lim_{n \to +\infty} \frac{x_n}{n} = g'(0) \quad \text{for almost every} \quad x \in \Omega . \tag{40}$$

In Ω we introduce a net $\{\mathfrak{M}_n\}$ as follows: For each $n = 0, 1, 2, \ldots$ $\mathfrak{M}_n$ consists of all sets of the form

$$\left\{x \in \Omega\colon\ x_{-n} \in \left[\frac{k_{-n}}{2^n}, \frac{k_{-n}+1}{2^n}\right), \ldots, x_n \in \left[\frac{k_n}{2^n}, \frac{k_n+1}{2^n}\right)\right\}, \quad k_{-n}, \ldots, k_0, \ldots, k_n \text{ integers.}$$

Obviously $\bigcup_{n=0}^{\infty} \mathfrak{M}_n$ generates $\mathfrak{B}_0 = \mathfrak{B}$.

Assume first that $g'(0) = \lambda_0$ is finite, let $\bar{x} = (\xi_n)$ be one of the sample points at which (40) holds and for each n let Y_n be the uniquely determined set in $\mathfrak{M}_n$ with $\bar{x} \in Y_n$. Y_n is of the form

$$Y_n = \{x\colon\ x_{-n} \in A_{-n}, \ldots, x_n \in A_n\}$$

where $A_{-n}, \ldots, A_n$ are dyadic intervals of length $1/2^n$.

Since $e^{s_0\lambda_0}/g(s_0) < 1$, there is a positive ε such that

$$\frac{e^{s_0(\lambda_0+\varepsilon)}}{g(s_0)} < 1 . \tag{41}$$

By (40) $\xi_n/n < \lambda_0 + \varepsilon$ when n is sufficiently large, i.e.

$$\xi_n < n(\lambda_0 + \varepsilon) .$$

Since the length of A_n is $1/2^n$ we have

$$x_n < n(\lambda_0 + \varepsilon) + \frac{1}{2^n} \quad \text{for all} \quad x_n \in A_n .$$

But then

$$e^{s_0 x_n} < e^{n s_0(\lambda_0+\varepsilon)+(s_0/2^n)}$$

and formula (33) yields

$$\frac{\mu_1(Y_n)}{\mu(Y_n)} \leqq \left[\frac{e^{s_0(\lambda_0+\varepsilon)+(s_0/n 2^n)}}{g(s_0)}\right]^n \quad \text{assuming} \quad \mu(Y_n) > 0 .$$

By (41)

$$\lim_{n\to+\infty} \frac{\mu_1(Y_n)}{\mu(Y_n)} = 0 .$$

If $g'(0)$ is infinite there are only two possibilities: Either $s_0 < 0$ and $g'(0) = +\infty$ or $s_0 > 0$ and $g'(0) = -\infty$. Assume the first and let $\tilde{x} = (\xi_n)$ be a sample point at which (40) holds. If $\{Y_n\}$ is defined as before, then

$$x_n < \xi_n + \frac{1}{2^n} \quad \text{for all} \quad x_n \in A_n ,$$

$$e^{s_0 x_n} < e^{s_0 \xi_n + (s_0/2^n)} \quad \text{for all} \quad x_n \in A_n$$

and by formula (33) again

$$\mu_1(Y_n) \leqq g(s_0)^{-n} e^{s_0\xi_n+(s_0/2^n)} \mu(Y_n) ,$$

$$\frac{\mu_1(Y_n)}{\mu(Y_n)} \leqq \left[\frac{e^{(s_0\xi_n/n)+(s_0/n 2^n)}}{g(s_0)}\right]^n \quad \text{if} \quad \mu(Y_n) > 0$$

and the right-hand side converges to 0 since $\xi_n/n \to +\infty$ and $s_0 < 0$.

We have thus shown that in all cases the Radon-Nikodym derivative of μ_1 with respect to μ is 0 a.e.

To handle μ_2 we can consider the reverse random walk, whose transition probabilities are given by

$$q(x, A) = p(-x, -A) .$$

The shift transformation of this random walk is isomorphic with the inverse T^{-1} of the shift of the original random walk. To show this it is sufficient to prove that $\text{Prob}\{x_k \in B \mid x_{k+1} = \xi\} = \text{Prob}\{x_{k+1} \in -B \mid x_k = -\xi\}$ i.e. for any Borel sets A, B in G

$$\int_A \lambda(d\xi)\, q(\xi, B) = \int_B \lambda(d\xi)\, p(\xi, A)$$

and this in turn follows easily from the translation invariance of λ since

$$\int_A \lambda(d\xi)\, q(\xi, B) = \cdots = \int_R p(0, d\eta) \int_{A\cap(\eta+B)} \lambda(d\xi) = \int_R p(0, d\eta) \int_{(A-\eta)\cap B} \lambda(d\xi)$$

$$= \cdots = \int_B \lambda(d\xi)\, p(\xi, A) .$$

The proof is complete.

Problem 8. If the state space is not a group and hence tools such as 5.4 are not available, how far can one go with a strong ratio limit property of the type of 5.1?

We wish to close this section with the remark that though the central limit theorem (with its error estimates) is not strong enough to imply quasi-mixing of random walks, nevertheless it does yield mixing in a special case. If the increment of the random walk is centred, non-lattice and has finite third moment and if for the sake of simplicity we assume that its variance is 1 then Theorem 2 in [9, § 42, p. 210] implies

$$\lim_{n\to+\infty} \sqrt{2\pi n}\, p^{(n)}(0, [a, b)) = b - a \tag{42}$$

uniformly in $-K \leqq a < b \leqq K$, and one can infer that for any bounded almost boundaryless subsets A, B of the real line with positive Lebesgue measure and any integer m

$$\lim_{n\to+\infty} \frac{p^{(n+m)}(x, A)}{p^{(n)}(y, B)} = \frac{\lambda(A)}{\lambda(B)}$$

uniformly in $-K \leqq x \leqq K$, $-K \leqq y \leqq K$. This is STONE's result for the process in question. From (42) we have the rate of convergence of $p^{(n)}(x, A)$ to 0 $(\sim n^{-1/2})$.

§ 6. Some Remarks on the Speed of Mixing and the Convergence Norm γ

As we have seen the simplest examples of quasi-mixing Markov chains at hand are the random walks (chains with independent increments) on the integers. It is known [36, p. 72] that for every such chain there is $A > 0$ such that

$$p_{00}^{(n)} \leqq \frac{A}{\sqrt{n}} \quad \text{for all } n$$

i.e. $\varrho_n \geqq B\sqrt{n}$. This means that the ergodic index of a random walk is at most 2. If the increment has finite variance, GNEDENKO's local limit theorem yields ([8, § 43, p. 297], [9, § 49, p. 233])

$$\lim_{n\to+\infty} \left[\sqrt{2\pi n}\, p_{0k}^{(n)} - \frac{1}{v}\exp\left(-\frac{n a^2}{2v^2} + k\frac{a}{v^2} - k^2\frac{1}{2nv^2}\right)\right] = 0 \tag{43}$$

uniformly in k (with a denoting the expectation and v^2 the variance of the increment). If $a = 0$ then by (43)

$$\lim_{n\to+\infty} \sqrt{2\pi n}\, p_{0k}^{(n)} = \frac{1}{v}$$

which is the discrete analogue of (42). If $a \neq 0$ then

$$\lim_{n\to+\infty} \sqrt{2\pi n}\, p_{0k}^{(n)} = 0$$

and if in addition we know that the increment has finite r-th moment ($r \geqq 3$) then a theorem of ESSEEN (see [9, Th. 1, § 51, p. 241]) implies

$$\lim_{n\to+\infty} n^{(r-1)/2} p_{0k}^{(n)} = 0\,.$$

The theorems quoted here have been generalized by HECKENDORFF ([12], [13], [14]) to finite dimensional random walks.

At the other end the centrally biased chains of GILLIS [7] (see also [17]) provide examples of mixing chains with much slower rates of convergence $\varrho_n \to +\infty$. In fact if in [17, § 3] we restrict ε to $(-\frac{1}{2}, \frac{1}{2}]$ then we have recurrence and an infinite positive stationary vector. The chains are then mixing (by reversibility

[30] or by the criterion of [25]) and as proved by GILLIS the chain determined by each ε has the following property: For any $\theta > 0$ there exists K_1 such that for all n

$$\frac{1}{K_1} n^{\varepsilon-\frac{1}{2}-\theta} < p_{00}^{(n)} < K_1 n^{\varepsilon-\frac{1}{2}+\theta}.$$

As shown in [17] this implies that if k is any positive integer and ε is in the interval

$$\left(\frac{1}{2} - \frac{1}{k}, \quad \frac{1}{2} - \frac{1}{k+1}\right)$$

then the ergodic index of the corresponding chain is k. For $\varepsilon = \frac{1}{2}$ we have a null-recurrent mixing chain with infinite ergodic index. (Trivially every positive-recurrent chain has infinite ergodic index.)

All of the quasi-mixing non-mixing examples we have encountered have convergence norm $\gamma < 1$, which means that $p_{00}^{(n)}$ converges to 0 geometrically and implies an extremely fast convergence of ϱ_n to $+\infty$, in particular much faster than for mixing chains. The question is open, however, of whether there exist a mixing chain and a quasi-mixing non-mixing chain such that the convergence $\varrho_n \to +\infty$ is faster for the mixing one. The quasi-mixing chain should obviously have $\gamma = 1$ and the mixing one should be transient since for recurrent chains

$$\sum_{n=1}^{\infty} p_{00}^{(n)} = +\infty,$$

hence $p_{00}^{(n)} \to 0$ slower. This leads us to:

Problem 9. Do all quasi-mixing non-mixing Markov chains have $\gamma < 1$? Here are two equivalent reformulations of this problem:

Problem 10. Does every quasi-mixing chain with $\gamma = 1$ have a unique (to within a constant factor) positive stationary vector?

Problem 11. Does $\gamma = 1$ imply that $\{\tau(i)\}$ is constant?

To be sure, if $\gamma = 1$ then (11) and (12) reduce to

$$\tau(i) = \sum_j p_{ij}\tau(j) \quad \text{and} \quad \pi(j) = \sum_i \pi(i) p_{ij}$$

so that the answer to Problem 9 is affirmative if and only if the answer to *both* Problems 10 and 11 are affirmative. However Problems 10 and 11 are equivalent, for if a quasi-mixing chain with $\gamma = 1$ has two distinct positive stationary vectors, $\{\pi(i)\}$ and say $\{\lambda_i\}$, then reversing such a chain with respect to $\{\lambda_i\}$ we shall find a new chain violating the implication in Problem 11 and *vice versa*.

In all of the concrete examples we know of, which have the SRLP and whose convergence norm is 1, the sequence $\{\tau(i)\}$ is constant. This is true of the chains with independent increments, of the model discussed in 2.5 and 2.6, or for instance the "balanced" chains of [20]. Another class of chains, where $\gamma = 1$ implies $\tau(i) =$ constant, are the "random walks" of [18] on the non-negative integers, which are reversible chains: Let $Q_i(x)$, $i = 0, 1, 2, \ldots$ be the polynomials of [18] and ψ the corresponding representing measure on $[-1, 1]$:

$$p_{ij}^{(n)} = \lambda_j \int_{-1}^{1} x^n Q_i(x) Q_j(x) d\psi(x).$$

Using the techniques of [18, § 3] one can show:

If

$$\lim_{n\to+\infty} \frac{p_{00}^{(n+1)}}{p_{00}^{(n)}}$$

exists, then the chain is quasi-mixing with

$$\lim_{n\to+\infty} \frac{p_{ij}^{(n+m)}}{p_{kh}^{(n)}} = \gamma^m \frac{Q_i(\gamma)\, Q_j(\gamma)\, \lambda_j}{Q_k(\gamma)\, Q_h(\gamma)\, \lambda_h} \tag{44}$$

where γ is the least positive number such that $[-\gamma, \gamma]$ contains the support of ψ. (44) is consistent with (17).

It was mentioned in § 2 that if a reversible chain is R-recurrent then it is quasi-mixing [34]. If it is R-transient $p_{00}^{(n+1)}/p_{00}^{(n)}$ may fail to converge. An example where $p_{00}^{(2n+1)}/p_{00}^{(2n)}$ and $p_{00}^{(2n+2)}/p_{00}^{(2n+1)}$ converge to distinct positive limits is given in [18, p. 77] (see also [34]). As observed by PRUITT however

$$\lim_{n\to+\infty} \frac{p_{00}^{(2n+2)}}{p_{00}^{(2n)}}$$

exists. In fact we can prove a little more.

Let (p_{ij}) be the transition matrix of an irreducible and reversible (but not necessarily aperiodic) Markov chain. D. G. KENDALL has shown ([21], [22]) that there are real signed finite measures μ_{ij} on $[-1, 1]$ such that

$$p_{ij}^{(n)} = \int_{-1}^{1} x^n \mu_{ij}(dx), \quad n = 0, 1, 2, \ldots. \tag{45}$$

When $i = j$ μ_{ij} is a positive measure. As before we denote by R the radius of convergence of

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} x^n$$

and by γ its reciprocal.

6.1. Theorem. *Let i be any state. Then γ is the least positive number such that the symmetric interval $[-\gamma, \gamma]$ contains the support of μ_{ii}. Further, for any integer m*

$$\lim_{n\to+\infty} \frac{p_{ii}^{(2n+m)}}{p_{ii}^{(2n)}} = \gamma^{2m}.$$

Proof. It is sufficient to demonstrate this for the state 0, since there is nothing special about this state. Choose any state i_1 such that $p_{0i_1} > 0$. By reversibility $p_{i_1 0} > 0$ hence $p_{00}^{(2)} \geqq p_{0i_1} p_{i_1 0} > 0$ and hence $p_{00}^{(2n)} > 0$ for every $n \geqq 1$.

The measure $\mu \equiv \mu_{00}$ is positive, therefore by (46) and the Schwarz inequality (see also [34])

$$(p_{00}^{(2n)})^2 = \left(\int_{-1}^{1} x^{2n} \mu(dx)\right)^2 = \left(\int_{-1}^{1} x^{n+1} x^{n-1} \mu(dx)\right)^2 \leqq$$

$$\leqq \left(\int_{-1}^{1} x^{2n+2} \mu(dx)\right)\left(\int_{-1}^{1} x^{2n-2} \mu(dx)\right) = p_{00}^{(2n+2)} p_{00}^{(2n-2)}$$

so that the sequence $p_{00}^{(2n+2)}/p_{00}^{(2n)}$ is increasing and

$$\lim_{n\to+\infty} \frac{p_{00}^{(2n+2)}}{p_{00}^{(2n)}} = \beta$$

exists. This implies

$$\lim_{n\to+\infty} \sqrt[n]{p_{00}^{(2n)}} = \beta .$$

On the other hand (cf. [24, Th. 10])

$$\lim_{n\to+\infty} \sqrt[n]{p_{00}^{(n)}}$$

exists (and is equal to γ of course). We conclude $\beta = \gamma^2$.

Now let α be the least positive number such that $[-\alpha, \alpha]$ contains the support of μ. Then

$$\gamma = \lim_{n\to+\infty} (p_{00}^{(2n)})^{1/2n} = \lim_{n\to+\infty} \left(\int_{-\alpha}^{\alpha} x^{2n} \mu(dx) \right)^{1/2n} \leqq \lim_{n\to+\infty} (\alpha^{2n} \mu([-\alpha, \alpha]))^{1/2n} = \alpha .$$

Conversely, given any $\varepsilon > 0$ denote by D_ε the union $[-\alpha, -\alpha+\varepsilon] \cup [\alpha-\varepsilon, \alpha]$. By the definition of α $\mu(D_\varepsilon) > 0$ for every $\varepsilon > 0$. Then

$$(p_{00}^{(2n)})^{1/2n} = \left(\int_{-\alpha}^{\alpha} x^{2n} \mu(dx) \right)^{1/2n} \geqq \left(\int_{D_\varepsilon} x^{2n} \mu(dx) \right)^{1/2n} \geqq [(\alpha-\varepsilon)^{2n} \mu(D_\varepsilon)]^{1/2n}$$
$$= (\alpha - \varepsilon)[\mu(D_\varepsilon)]^{1/2n} .$$

Then

$$\gamma = \lim_{n\to+\infty} (p_{00}^{(2n)})^{1/2n} \geqq \alpha - \varepsilon$$

and since this is true for every $\varepsilon > 0$ we deduce $\gamma \geqq \alpha$.

One might perhaps think that the supports of all the measures μ_{ij} are contained in $[-\gamma, \gamma]$. This would imply that γ is the norm of the linear transformation T defined in [21], in terms of which the integral representation is achieved. (Recall that if the chain is reversible this transformation is self-adjoint, so that no dilation is required for the representation.) This is not always true however. Consider the random walk on the integers with

$$p_{i,i-1} = q^2, \quad p_{ii} = 2pq, \quad p_{i,i+1} = p^2$$

where $p \neq q$, $p + q = 1$, $p > 0$. Here $\gamma < 1$ by 4.5. If however we define

$$\omega_0 = (\ldots, 0, 0, 0, 1, 0, 0, 0, \ldots),$$
$$\omega_1 = (\ldots, 0, 0, 1, 1, 1, 0, 0, \ldots),$$
$$\omega_2 = (\ldots, 0, 1, 1, 1, 1, 1, 0, \ldots),$$
etc.

then $\{\omega_n\}$ is a sequence of elements of the Hilbert space l^2 such that

$$\lim_{n\to+\infty} \frac{\| T\omega_n \|}{\| \omega_n \|} = 1$$

where T denotes KENDALL's linear transformation.

§ 7. Some Pathological Examples of Transformations

In this section we return to the general theme of § 1 to construct a couple of "pathological" examples and offer further justification for our policy of accepting only products of the form $\mu_1(E)\mu_2(F)$ in (3). In [28] the right-hand side of (3) is allowed to be any function $\varphi(E, F)$ on pairs of sets, which is a σ-finite tight measure in each variable if the other one is fixed and which has the following property: If $\mu(E) > 0$, $\mu(F) > 0$ and E, F are μ-almost boundaryless, then $\varphi(E, F) > 0$.

7.1 Propositions. *There exists a transformation on the unit interval* $[0, 1]$ *which satisfies Krickeberg's original definition of quasi-mixing but is not mixing in the classical sense (not even ergodic) in* $[0, 1]$.

Let μ stand for the Lebesgue measure in $[0, 1]$.

7.2. Lemma. *Let K be a nowhere dense subset of* $[0, 1]$ *with* $\mu(K) > 0$. *If T is an invertible measure-preserving transformation which is mixing in the classical sense on* $L = [0, 1] - K$ *and equal to the identity transformation on K (i.e.* $Tx = x$ *for every* $x \in K$*) then for all measurable sets A, B*

$$\lim_{n\to+\infty} \mu(A \cap T^{-n} B) = \mu(A \cap B \cap K) + \frac{\mu(A \cap L)\,\mu(B \cap L)}{\mu(L)}\,.$$

where the right-hand side is positive whenever A, B are μ-almost boundaryless and have positive measure.

In fact, if A, B are almost boundaryless with $\mu(A) > 0$, $\mu(B) > 0$, then $\mu(A - K) > 0$, $\mu(B - K) > 0$. For if we assume $\mu(A - K) = 0$ we should have $\mu(A^0 - \bar{K}) = 0$, $A^0 - \bar{K} = \emptyset$ since $A^0 - \bar{K}$ is open. But then $A^0 \subset \bar{K}$, which implies $A^0 = \emptyset$ since K is nowhere dense. This leads to the contradiction $\mu(A) = \mu(A^0) = 0$.

There remains the more difficult task of constructing such a transformation with the additional property of almost everywhere continuity. Let $0 < \alpha < 1$ and from the interval $[0, 1]$ delete the closed interval $\Delta_{0,1}$ of length $\alpha/2$, occupying middle position in $[0, 1]$. From the two remaining intervals delete the middle closed subintervals $\Delta_{1,1}$, $\Delta_{1,2}$ of length $\alpha/8 = \alpha/2^3$, then subintervals $\Delta_{2,1}$, $\Delta_{2,2}$, $\Delta_{2,3}$, $\Delta_{2,4}$ of length $\alpha/32 = \alpha/2^5$ etc. Proceeding as in the construction of the Cantor set we delete at the n-th step intervals $\Delta_{n,1}, \Delta_{n,2}, \ldots, \Delta_{n,2n}$ each of length $\alpha/2^{2n+1}$. Let

$$L = \bigcup_{n,i} \Delta_{n,i}$$

and $K = [0, 1] - L$. The set K is nowhere dense, $\mu(K) = 1 - \alpha$, $\mu(L) = \alpha$.

We shall define the transformation T on L as a "Markov chain" on the intervals $\Delta_{n,i}$. More precisely we conceive of $\{\Delta_{n,i}\}$ as the states of a suitably defined Markov chain and then use KRICKEBERG's isomorphism theorem [28, § 2] to get a transformation on L. The transition probabilities of our chain will be the following:

For $n \geqq 1$ and $i = 2k - 1$ odd:

$$\operatorname{Prob}[\Delta_{n,2k-1} \to \Delta_{n-1,k}) = \tfrac{1}{2}\,,$$
$$\operatorname{Prob}[\Delta_{n,2k-1} \to \Delta_{n,2k-1}] = \tfrac{1}{2}\,.$$

For $n \geqq 1$ and $i = 2k$ even:

$$\mathrm{Prob}[\Delta_{n,2k} \to \Delta_{n-1,k}] = \tfrac{1}{2},$$
$$\mathrm{Prob}[\Delta_{n,2k} \to \Delta_{n,2k}] = \tfrac{1}{2}.$$

For the state $\Delta_{0,1}$

$$\mathrm{Prob}[\Delta_{0,1} \to \Delta_{n,i}] = \frac{1}{2^{2(n+1)}}, \quad n \geqq 1;\ i = 1, 2, \ldots, 2^n,$$
$$\mathrm{Prob}[\Delta_{0,1} \to \Delta_{0,1}] = \tfrac{3}{4}.$$

This Markov chain is irreducible and aperiodic and preserves the measure

$$\lambda(\Delta_{n,i}) = \alpha/2^{2n+1}$$

which is the length of the interval $\Delta_{n,i}$. Since

$$\sum_{n,i} \lambda(\Delta_{n,i}) = \alpha$$

is finite, the chain is mixing (2.2). We then "translate" the shift transformation of this chain into a transformation T on L by KRICKEBERG's method. To be sure this construction yields an isomorphism of the sample space $(\Omega, \mathfrak{B}, \mu)$ of the above chain and a space $(E, \mathfrak{F}, \nu)$ where E is a union

$$\bigcup_{n,i} \Delta'_{n,i}$$

of rectangles in the plane and ν the two-dimensional Lebesgue measure, but we can easily modify the construction (for instance by subdividing the intervals $\Delta_{n,i}$ *alternately* according to "future" and "past" time instants $n = 0, 1, -1, 2, -2, \ldots$) to get an isomorphism S of Ω onto L which carries over the shift of Ω to the required transformation T. The latter is almost everywhere continuous on L.

If we define $Tx = x$ on K the extension $T\,|\,[0, 1]$ is continuous at every point of K. In fact, note that if $x \in \Delta_{n,2k-1}$ then $Tx \in \Delta_{n,2k-1} \cup \Delta_{n-1,k}$ and if $x \in \Delta_{n,2k}$ then $Tx \in \Delta_{n,2k} \cup \Delta_{n-1,k}$. This clearly implies that if $x_m \to x$, where $x_m \in L$, $x \in K$, then $|x_m - Tx_m| \to 0$ so that $Tx_m \to Tx$.

Note that in this example we can choose α as small as we wish. The transformation T will then be completely "calm" on K (with $\mu(K) = 1 - \alpha$) and the "stirring" will occur only in a set of small measure.

A modification of the above example yields a similar transformation on $[0, +\infty)$. Consider the half-line

$$[0, +\infty) = \bigcup_{\nu=0}^{\infty} [\nu, \nu + 1];$$

in each $[\nu, \nu + 1]$ construct a set

$$L_\nu = \bigcup_{n,i} \Delta_{n,i}^{(\nu)}$$

as before, with measure $\mu(L_\nu) = \alpha_\nu = 1/2^{\nu+1}$ and define a Markov chain as follows. In each L_ν the transitions will be as before, with the following exception: Where we had

$$\mathrm{Prob}[\Delta_{0,1} \to \Delta_{0,1}] = \tfrac{3}{4}$$

before, we now have

$$\text{Prob}[\Delta_{0,1}^{(\nu)} \to \Delta_{0,1}^{(\nu+1)}] = \tfrac{3}{8},$$
$$\text{Prob}[\Delta_{0,1}^{(\nu)} \to \Delta_{0,1}^{(0)}] = \tfrac{3}{8}.$$

We continue as before and get a "quasi-mixing" (in the sense of KRICKEBERG) transformation T on $[0, +\infty)$ having $\varrho_n = 1$ for every n. In fact for all integrable sets A, B

$$\lim_{n\to+\infty} \mu(A \cap T^{-n} B) = \mu(A \cap B \cap K) + \frac{\mu(A \cap L)\,\mu(B \cap L)}{\mu(L)}$$

where
$$L = \bigcup_{\nu=0}^{\infty} L_\nu, \quad K = [0, +\infty) - L, \quad \mu(L) = 1.$$

Theorem 1.2 shows that by requiring that the right-hand side of (3) should have the form $\mu_1(E)\,\mu_2(F)$ we eliminate the pathology of the first of the above examples. However the pathology of the second example may still creep up:

7.3. Proposition. *There exists a quasi-mixing (in the sense of Definition* 1.1*) transformation on a space of infinite measure, such that* $\varrho_n = 1$ *for every* n.

Given a space X with $\mu(X) = +\infty$, choose a subset L of finite measure, let T be mixing on L (in the classical sense) and such that on the complement K of L it satisfies

$$\lim_{n\to+\infty} \mu(A \cap T^{-n} B) = 0$$

whenever $\mu(A) < +\infty$, $\mu(B) < +\infty$, $A \subset K$, $B \subset K$. The following concrete construction will render K nowhere dense and T quasi-mixing and almost everywhere continuous.

Let R be the real line,

$$R = \bigcup_{\nu=-\infty}^{\infty} [\nu, \nu + 1].$$

Fix $0 < \alpha < 1$ and in each $[\nu, \nu + 1]$ make the construction of the first example. Denote by $\Delta_{n,i}^{(\nu)}$ the corresponding intervals. Then let

$$L_\nu = \bigcup_{n \geqq |\nu|} \bigcup_{i=1}^{2^n} \Delta_{n,i}^{(\nu)},$$
$$L = \bigcup_{\nu=-\infty}^{+\infty} L_\nu, \quad K = R - \bigcup_{\text{all } \nu,n,i} \Delta_{n,i}^{(\nu)}.$$

Our space X will be $L \cup K$. Clearly $\mu(L) = 3\alpha$. On K we define $Tx = x + 1$. On L T is determined by means of a Markov chain as follows:

Case (i) $\nu \geqq 1$. If $n \geqq \nu + 2$, then

$$\text{Prob}[\Delta_{n,2k-1}^{(\nu)} \to \Delta_{n-1,k}^{(\nu+1)}] = 1,$$
$$\text{Prob}[\Delta_{n,2k}^{(\nu)} \to \Delta_{n-1,k}^{(\nu+1)}] = 1.$$

Transitions from $\Delta_{\nu+1,i}^{(\nu)}$ $(i = 1, 2, \ldots, 2^{\nu+1})$ are as follows

$$\text{Prob}[\Delta_{\nu+1,i}^{(\nu)} \to \Delta_{n,j}^{(\nu+1)}] = 2^{\nu+1}/2^{2n+1}, \quad n \geqq \nu + 1; \quad \text{any } i, j.$$

Transitions from $\Delta_{\nu,i}^{(\nu)}$ $(i = 1, 2, \ldots, 2^\nu)$ are given by

$$\mathrm{Prob}[\Delta_{\nu,i}^{(\nu)} \to \Delta_{n,j}^{(0)}] = 1/2^{2n+1} \quad \text{any} \quad n, i, j\,.$$

Case (ii). $\nu \leqq -1$

$$\left.\begin{aligned} &\mathrm{Prob}[\Delta_{n,2k-1}^{(\nu)} \to \Delta_{n-1,k}^{(\nu+1)}] = 1\,, \\ &\mathrm{Prob}[\Delta_{n,2k}^{(\nu)} \to \Delta_{n-1,k}^{(\nu+1)}] = 1 \end{aligned}\right\} \quad \text{any} \quad n\,(\geqq |\nu|)\,.$$

Case (iii). $\nu = 0$. If $n \geqq 2$, then

$$\mathrm{Prob}[\Delta_{n,2k-1}^{(0)} \to \Delta_{n-1,k}^{(1)}] = 1\,,$$
$$\mathrm{Prob}[\Delta_{n,2k}^{(0)} \to \Delta_{n-1,k}^{(1)}] = 1\,.$$

Transitions from $\Delta_{1,i}^{(0)}$ $(i = 1, 2)$ are given by

$$\mathrm{Prob}[\Delta_{1,i}^{(0)} \to \Delta_{n,j}^{(1)}] = 1/2^{2n}, \quad n \geqq 1; \quad \text{any } i, j\,.$$

Transitions from $\Delta_{0,1}^{(0)}$

$$\mathrm{Prob}[\Delta_{0,1}^{(0)} \to \Delta_{n,j}^{(\nu)}] = 1/2^{2n+1}, \quad \nu \leqq -1; \quad n \geqq |\nu|; \quad \text{any } j\,.$$

This chain is irreducible and aperiodic and preserves the measure (length) of each $\Delta_{n,i}^{(\nu)}$. Since $\mu(L) = 3\alpha < +\infty$ the chain is mixing. We can show as in the first example that the transformation T on $X = L \cup K$ thus defined is almost everywhere continuous and that for all integrable subsets A, B of X

$$\lim_{n\to+\infty} \mu(A \cap T^{-n} B) = \frac{\mu(A \cap L)\, \mu(B \cap L)}{\mu(L)}\,.$$

The right-hand side is of the form $\mu_1(A)\, \mu_2(B)$ and $\varrho_n = 1$ for every n.

We conclude our pathological examples with:

7.4. Proposition. *There exists a quasi-mixing transformation operating on a space with atoms.*

Let for instance R be the real line and for every subset $A \subset R$ denote by A^* the set one gets from A by deleting all integers in A. The measure μ will be the Lebesgue measure on R^* and will assign mass 1 to each integer. Let T be any transformation on R^* which is mixing for bounded (in the ordinary sense) almost boundaryless sets. We extend T to R by defining $Tn = n + 1$. It is easy to see that for any two bounded and almost boundaryless sets A, B in R we have $A \cap T^{-n} B = A^* \cap T^{-n} B^*$ for sufficiently large n, hence if $\{\varrho_n\}$ is the sequence governing the mixing behaviour of T on R^* then

$$\lim_{n\to+\infty} \varrho_n\, \mu(A \cap T^{-n} B) = \mu(A^*)\, \mu(B^*) = \mu_1(A)\, \mu_1(B)$$

where μ_1 is the Lebesgue measure on R.

Further, if E is an almost boundaryless (relative to μ) set with $\mu(E) > 0$, then any integers contained in E must be interior points of E so that $\mu_1(E) > 0$. This takes care of condition (ii) in Definition 1.1.

References

1. Blackwell, D.: On transient Markov processes with a countable number of states and stationary transition probabilities. Ann. math. Statistics **26**, 654—658 (1955).
2. Chung, K. L.: Markov chains with stationary transition probabilities. Berlin-Göttingen-Heidelberg: Springer 1960.
3. —, and P. Erdös: Probability limit theorems assuming only the first moment. Mem. Amer. math. Soc. **6**, 1951.
4. —, and W. H. J. Fuchs: On the distribution of values of sums of random variables. Mem. Amer. math. Soc. **6**, 1951.
5. Derman, C.: Some contributions to the theory of denumerable Markov chains. Trans. Amer. math. Soc. **79**, 541—555 (1955).
6. Doob, J. L.: Stochastic processes. New York-London-Sydney: Wiley 1953.
7. Gillis, J.: Centrally biased discrete random walk. Quart. J. Math., Oxford II. Ser. **7**, 144—152 (1956).
8. Gnedenko, B. V.: The theory of probability, 2nd edition. New York: Chelsea 1962.
9. —, and A. N. Kolmogorov: Limit distributions for sums of independent random variables. Reading-Palo Alto-London: Addison-Wesley 1954.
10. Hajian, A. B., and S. Kakutani: Weakly wandering sets and invariant measures. Trans. Amer. math. Soc. **110**, 136—151 (1964).
11. Halmos, P. R.: Lectures on ergodic theory. New York: Chelsea 1956.
12. Heckendorff, H.: On a multi-dimensional local limit theorem for lattice distributions (in Ukrainian). Dopovidi Akad. Nauk ukrain. RSR 19—23, **1965**.
13. — Multi-dimensional local limit theorem for densities (in Russian). Ukrain. mat. Žurn. **16**, 365—373 (1964).
14. — On a multi-dimensional local limit theorem for distributions of mixed type (in Russian). First Republ. Math. Conf. of Young Researchers, Part II, pp. 690—697. Akad. Nauk ukrain. SSR, Inst. Mat., Kiev 1965.
15. Hewitt, E., and K. A. Ross: Abstract harmonic analysis, Vol. 1. Berlin-Göttingen-Heidelberg: Springer 1963.
16. Hopf, E.: Ergodentheorie. Berlin: Springer 1937 (reprinted: New York: Chelsea 1948).
17. Kakutani, S., and W. Parry: Infinite measure preserving transformations with "mixing". Bull. Amer. math Soc. **69**, 752—756 (1963).
18. Karlin, S., and J. McGregor: Random walks. Illinois J. Math. **3**, 66—81 (1959).
19. Kemeny, J. G.: A probability limit theorem requiring no moments. Proc. Amer. math. Soc. **10**, 607—612 (1959).
20. — Representation theory for denumerable Markov chains. Trans. Amer. math Soc. **125**, 47—62 (1966).
21. Kendall, D. G.: Unitary dilations of Markov transition operators and the corresponding integral representation for transition-probability matrices. Harold Cramér Volume, pp. 139—161. Stockholm: Almquist and Wiksell 1960.
22. — Integral representations for Markov transition probabilities. Bull. Amer. math. Soc. **64**, 358—362 (1958).
23. Kesten, H.: Ratio theorems for random walks II. J. d'Analyse math. **11**, 323—379 (1963).
24. Kingman, J. F. C.: The exponential decay of Markov transition probabilities. Proc. London math. Soc., III. Ser. **13**, 337—358 (1963).
25. —, and S. Orey: Ratio limit theorems for Markov chains. Proc. Amer. math. Soc. **15**, 907—910 (1964).
26. Krengel, U.: On the global limit behaviour of Markov chains and of general non-singular Markov processes. Z. Wahrscheinlichkeitstheorie verw. Geb. **6**, 302—316 (1966).
27. Krickeberg, K.: Probability theory. New York: Addison-Wesley 1965.
28. — Strong mixing properties of Markov chains with infinite invariant measure. Proc. Fifth Berkeley Sympos. math. Statist. Probab., Vol II, Part 2, pp. 431—446. Berkeley and Los Angeles: University of California Press 1967.
29. Neveu, J.: Bases mathématiques du calcul des probabilités. Paris: Masson 1964.
30. Orey, S.: Strong ratio limit property. Bull. Amer. math. Soc. **67**, 571—574 (1961).

31. ORNSTEIN, D.: A limit theorem for independent random variables, unpublished.
32. PAPANGELOU, F.: On Markov processes whose shift transformation is quasi-mixing. Sympos. Probability Methods in Analysis, Lecture Notes in Mathematics No. 31, pp. 272–279. Berlin-Heidelberg-New York: Springer 1967.
33. — Footnote to the construction of σ-finite measures in product spaces, to appear.
34. PRUITT, W. E.: Strong ratio limit property for R-recurrent Markov chains. Proc. Amer. math. Soc. **16**, 196–200 (1965).
35. SAKS, S.: Theory of the integral. Warszawa-Lwow 1937.
36. SPITZER, F.: Principles of random walk. Princeton — New York: van Nostrand 1964.
37. STONE, C.: Ratio limit theorems for random walks on groups. Trans. Amer. math. Soc. **125**, 86–100 (1966).
38. VERE-JONES, D.: Geometric ergodicity in denumerable Markov chains. Quart. J. Math. (Oxford), II. Ser. **13**, 7–28 (1962).

Dr. F. Papangelou
Institut f. Angew. Mathematik
der Universität
6900 Heidelberg, Berliner Straße 17